EPA Public Access

Author manuscript

J Pharmacokinet Pharmacodyn. Author manuscript; available in PMC 2018 December 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

J Pharmacokinet Pharmacodyn. 2017 December ; 44(6): 549–565. doi:10.1007/s10928-017-9548-7.

Evaluation and calibration of high-throughput predictions of chemical distribution to tissues

Robert Pearce^{1,2}, R. Woodrow Setzer¹, Jimena Davis^{1,§}, and John F. Wambaugh^{1,*}

¹National Center for Computational Toxicology, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, North Carolina 27711

²Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee 37831

Abstract

Toxicokinetics (TK) provides critical information for integrating chemical toxicity and exposure assessments in order to determine potential chemical risk (*i.e.*, the margin between toxic doses and plausible exposures). For thousands of chemicals that are present in our environment, *in vivo* TK data are lacking. The publicly available R package “httk” (version 1.8, named for “high throughput TK”) draws from a database of *in vitro* data and physico-chemical properties in order to run physiologically-based TK (PBTK) models for 553 compounds. The PBTK model parameters include tissue:plasma partition coefficients (K_p) which the httk software predicts using the model of Schmitt (Toxicol In Vitro, 2008). In this paper we evaluated and modified httk predictions, and quantified confidence using *in vivo* literature data. We used 964 rat K_p measured by *in vivo* experiments for 143 compounds. Initially, predicted K_p were significantly larger than measured K_p for many lipophilic compounds ($\log_{10}$ octanol:water partition coefficient > 3). Hence the approach for predicting K_p was revised to account for possible deficiencies in the *in vitro* protein binding assay, and the method for predicting membrane affinity was revised. These changes yielded improvements ranging from a factor of 10 to nearly a factor of 10000 for 83 K_p across 23 compounds with only 3 K_p worsening by more than a factor of 10. The vast majority (92%) of K_p were predicted within a factor of 10 of the measured value (overall root mean squared error of 0.59 on $\log_{10}$ -transformed scale). After applying the adjustments, regressions were performed to calibrate and evaluate the predictions for 12 tissues. Predictions for some tissues (e.g., spleen, bone, gut, lung) were observed to be better than predictions for other tissues (e.g., skin, brain, fat), indicating that confidence in the application of *in silico* tools to predict chemical partitioning varies depending upon the tissues involved. Our calibrated model was then evaluated using a second data set of human *in vivo* measurements of volume of distribution (V_{ss}) for 498 compounds reviewed by Obach et al. (2008). We found that calibration of the model improved performance: a regression of the measured values as a function of the predictions has a slope of 1.03, intercept of -0.04, and R^2 of 0.43. Through careful evaluation of predictive methods for

*Corresponding Author: John F. Wambaugh, 109 T.W. Alexander Dr, NC 27711, USA, Wambaugh.John@epa.gov, Phone: (919) 541-7641.

§Current Address: Jimena Davis is now employed by Syngenta. Dr. Davis received no funding from Syngenta for this project, and Syngenta was not involved in the development or approval of this research or this report.

Disclaimer: The United States Environmental Protection Agency, through its Office of Research and Development, funded and managed the research described here. However, it may not necessarily reflect official Agency policy, and reference to commercial products or services does not constitute endorsement.

chemical partitioning into tissues, we have improved and calibrated these methods and quantified confidence for TK predictions in humans and rats.

Introduction

Analysis of pooled human plasma samples has indicated the presence, mostly in low levels, of more than 2000 potentially anthropogenic compounds [1]. Many more thousands of chemicals have been indicated in a recent analysis of house dust samples from homes in the United States [2]. The presence of these chemicals in our blood and our environment may be due to widespread commercial use of as many as 30,000 different chemicals [3]. Ideally these thousands of chemicals would be screened on the basis of risk posed to human health.

Two key factors that drive chemical risk assessment are toxicity and exposure potential [4]. However, toxicokinetic (TK) models are needed to predict tissue concentrations due to exposure in order to estimate the margin between toxic doses and exposures [5,6]. Traditional TK methods are resource intensive, often requiring hundreds of test animals per chemical [7] and either the development of radiolabeled compound [8] or analytical chemistry methods of sufficient quality to quantify chemical concentration in tissue [9]. Most of these chemicals are non-therapeutic (e.g., industrial plasticizers) and therefore both ethical and cost considerations limit human TK testing [10]. For the many thousands of chemicals that need TK, even animal testing is economically out of reach [7]. As an alternative, relatively high throughput TK (HTTK) methods have been developed, based upon chemical structure-derived properties and limited *in vitro* assays that allow tissue concentrations to be predicted without the need for precise *in vivo* data [11,12,6]. HTTK methods are used by the pharmaceutical industry to determine the range of efficacious doses and to prospectively evaluate success of planned human clinical trials [13–15].

Physiologically based toxicokinetic (PBTK) models vary in complexity and certainty depending upon the available data for estimating model parameters. PBTK models present several advantages over more empirical models (e.g., the two compartment model [16]): mapping the mechanics of the models onto physiologic processes provides an interpretation for parameter values that can speed their estimation by constraining the plausible values [17,18]. PBTK models also facilitate interspecies extrapolation since some processes (driven by chemical properties), including partitioning of chemical into various cellular constituents (e.g., lipid), are reasonably assumed to be consistent across species even as the blood flows, volumes of tissues, and fractional volumes of tissue constituents vary in known and well characterized ways [19–21].

Partition coefficients, defined as the concentration ratio of two adjacent compartments at equilibrium, are used in some PBTK models in determining individual tissue concentrations of compounds relative to blood or plasma. Many models with similar structures have been developed to determine tissue: plasma partition coefficients (K_p) [22–26]. However, high-throughput screening for thousands of chemicals requires a model that can be parameterized for many compounds using a small set of easily generated data. Schmitt's model [22] predicts red blood cell to plasma partitioning, and thus blood to plasma partition coefficients [27], and requires knowledge of only lipophilicity (*i.e.*, P_{ow}), the fraction unbound in plasma

(f_{up}), and chemical ionization equilibria (pK_a). These parameters can be predicted with quantitative structure-activity relationships (QSAR) when measurements are unavailable. The f_{up} is typically measured *in vitro*, though models for predicting f_{up} from chemical structure have recently become available [28,29].

The K_p are also used in calculating steady-state volume of distribution, V_{ss} , defined as the implied volume (*e.g.*, L) that yields the peak concentration (*e.g.*, mg/L) as observed in plasma given an intravenous dose (*e.g.*, mg). The V_{ss} is calculated by summing all K_p , each multiplied by its respective tissue volume, and the plasma volume [22]. Multiplying V_{ss} by the plasma concentration yields the total amount of a compound within a person or animal, assuming equilibrium. The V_{ss} determines the initial concentration and elimination rate of one compartment models, and can be measured *in vivo* from blood samples, allowing further K_p model evaluation in a human context [30].

In this report we compare predictions of K_p and V_{ss} from Schmitt's model (2008) (with modifications) to *in vivo* data in order to guide revision and calibration. These improved and calibrated predictors then can be used to predict chemical partitioning in humans for thousands of chemicals that are unlikely to ever have a human clinical trial. Estimated standard errors establish tissue-specific levels of confidence for the prediction of chemical partitioning for use in risk-based decision making.

Methods

We collected 964 rat K_p values estimated from *in vivo* experiments for 143 compounds and used them to evaluate and calibrate the predictions of Schmitt's model [22]. First, we implemented refinements to the Schmitt (2008) model that improve the agreement between predicted and *in vivo* measured K_p values. Next, we quantified the remaining differences between *in silico* predictions and *in vivo* measurements via linear regressions of the *in vivo* data on our predicted values to estimate calibration coefficients and standard errors for K_p predictions. We evaluated these calibrated predictions by using them to predict V_{ss} and blood to plasma partitioning for humans and compared them against measured data for 498 and 279 chemicals, respectively [30]. We provide our revised algorithm with the regressions in the freely available, open source R software package "httk".

The K_p and V_{ss} were initially predicted with the method contained in R package httk version 1.5 [31–34] which uses the model described in Schmitt [22] with an additional regression for calculating membrane affinity [10]. We note that, in this methods section, K with the upper-case subscript "P" refers to protein:water partition coefficient (K_p), while throughout the rest of the text K with the lower-case subscript "p" refers to tissue:plasma partition coefficients (K_p). The revised methods and data described here were released with R package "httk" version 1.8, which includes additional values for species-specific tissue and plasma component fractional volumes, P_{ow} , pK_a , and f_{up} that both provide data for new compounds and revise data for compounds previously contained in the package. Throughout our evaluations we use the root-mean square errors (RMSE) between the $\log_{10}$ predicted and measured values as a guide to overall predictive ability, with lesser RMSE being preferred. We note that the difference between the logarithms of two numbers is equivalent to the

logarithm of the ratio of the two number, and therefore the RMSE is a measure of the bias and variance of the absolute fold-change (up or down).

Data

Sources of Measurements of K_p

The evaluation used 964 rat K_p measurements of 143 compounds from 11 sources [22,23,25,35–42] that included both reviews and primary papers. The K_p measurements for a given compound and tissue sometimes varied across sources, with values reported as an individual measurement, a range of values, or the mean of multiple measurements. For each K_p , duplicate values were removed, and the means of the individual values were used. Values from Yun, Edginton [41] and Rodgers et al. [36] were used when it was not clear if there were duplicates within the means. The K_p were used from Gueorguieva et al. [38] that were calculated with the closed loop method. The K_p were measured *in vivo* with a variety of dosing methods including intravenous infusion and bolus injection. Rat red blood cell K_p were calculated from *in vitro* blood to plasma partitioning (B:P ratio), using a hematocrit value of 0.46 [43], assuming the remaining blood volume to be plasma. Human B:P ratio, used in a separate evaluation, were taken from 6 sources [10,12,44–47] and converted into red blood cell K_p in the same way as for rats but with a hematocrit value of 0.44 [43].

Sources of Measurements of pK_a , f_{up} , P_{ow} , tissue volumes, and V_{ss}

The pK_a and f_{up} values were taken from their respective K_p data sources, and when unavailable, pK_a was calculated with the ChemAxon pK_a prediction [48,49]. The P_{ow} were also taken from their respective K_p data source where available. The sources for the compound specific data ($\log P$, membrane affinity, f_{up} , and pK_a) used for rat data are included in Supplemental Table 1, and the sources used with human data are included in httx version 1.8. Species-specific (rat and human) fractional tissue component and plasma protein volumes came from Ruark et al. [27], replacing the Schmitt [22] values for all tissues except for bone and gut. Similar tissue volume data were used in Schmitt [22] as in Ruark et al. [27] for rats, with a few differences. However, the use of the latter data yielded significantly better results and uncovered an error in the fractional volume of neutral lipid in the brain in Schmitt's tissue data set, which had led to a systematic over-prediction of K_p in brain. Thirteen compounds were excluded because of their arbitrary f_{up} value of 0.5. All other rat K_p data available for the evaluated tissues were used except for Laniquidar, which had a f_{up} of 0, and brain K_p of Sparfloxacin, also 0. In the adjustment of f_{up} , the fractional volumes of neutral lipid and neutral phospholipid were taken from Peyret et al. [26] for rats and Poulin, Haddad [50] for humans. Empirical data for 498 compounds was taken from Obach et al. [30] to evaluate V_{ss} . These V_{ss} were inferred from a variety of *in vivo* studies administering i.v. dosing to humans [29]. Tissue volumes, L/kg, for calculating V_{ss} were calculated by converting the fractional mass of each tissue with its density [51,52], lumping the remaining tissues into the rest-of-body, excluding the mass of the gastrointestinal contents. Hematocrit was taken from Davies, Morris [43]. Membrane affinity values, used in K_p calculation when available, were taken from Schmitt [22], Endo et al. [53], and other sources cited in the package.

Schmitt's Model

In Schmitt's model, each tissue is composed of cells and interstitium, with each cell consisting of neutral lipid, neutral phospholipid, water, protein, and acidic phospholipid. Each tissue cell is defined as the sum of separate compartments for each constituent, all of which partition with a shared water compartment. The partitioning between the cell components and cell water is compound specific and determined by P_{ow} (in neutral lipid partitioning), membrane affinity (phospholipid and protein partitioning), and pK_a (neutral lipid and acidic phospholipid partitioning). For a given compound the partitioning into each component is identical across tissues. Thus the differences among tissues are driven by their composition, i.e. the varying volumes of components such as neutral lipid. However, pH differences across tissues also determine small differences in partitioning between cell and plasma water. The f_{up} is used as the plasma water to total plasma partition coefficient and to approximate the partitioning between interstitial protein and water.

The partition coefficients for all components of a cell are summed together to calculate the cell:cell water partition coefficient [22]:

$$K_{cell} = F_W + F_{nL}K_{nL} + F_{nPL}K_{nPL} + F_{aPL}K_{aPL} + F_PK_P \quad (1)$$

where $F_{component}$ represents the fractional volumes of the cell and $K_{component}$ represents the component:cellular water partition coefficients for water (W), neutral lipid (nL), neutral phospholipid (nPL), acidic phospholipid (aPL), and protein (P).

The total amount in the cell, A_{cell} , is defined as the sum of the amounts across all cell compartments in the model. V_{cell} , the total cell volume, can be larger than the total volume of the model components, assuming negligible partitioning into components beyond the scope of the model, but except for in the case of bone and gut, it is taken to be the sum of only the model components. The concentration in the cell is given by:

$$C_{cell} = \frac{A_{cell}}{V_{cell}} = \sum_{components} \frac{V_{component}C_{component}}{V_{cell}} = \sum_{components} F_{component}C_{component} \quad (2)$$

K_{tissue} is defined as the partition coefficient between the tissue components and the plasma water:

$$K_{tissue} = F_{cell}K_{cellular\ water:plasma\ water}K_{cell} + F_{int}K_{int} \quad (3)$$

where we sum the components of the tissue, cell and interstitium, together in the same way as the components of the cell [22]. Since interstitial water is assumed to have the same pH as plasma water, $K_{interstitial\ water:plasma\ water} = 1$, and this factor can be ignored. However, to account for the concentration difference between the cellular water and plasma water for each tissue due to pH differences, K_{cell} is converted with $K_{cellular\ water:plasma\ water}$ [22]:

$$K_{\text{cell water:plasma water}} = \frac{D_{\text{plasma}}}{D_{\text{cell}}} \quad (4)$$

where D represents the distribution coefficient:

$$D = K_{nL} = P_{ow}(F_{\text{neutral}} + \alpha F_{\text{charged}}) \quad (5)$$

with $\alpha = 0.001$ [22]. Equation 5, assumes that the partitioning into neutral lipid, D_{cell} , is proportional to the rate of the compound exiting either compartment through the cell membrane ($D_x C_x$), resulting in an equilibrium concentration in each compartment proportional to the ionized fraction of the compound. D_{cell} is also used as the neutral lipid partition coefficient of the tissue, K_{nL} . Ionization is calculated with the Henderson-Hasselbalch equation using pH values from Schmitt [22] for the cells, and a pH of 7.4 for plasma. The neutral fraction of zwitterions is treated as neutral.

f_{up} is used as $K_{\text{plasma water:plasma}}$ in converting K_{tissue} to K_p [22]:

$$K_p = f_{up} K_{\text{tissue}} \quad (6)$$

The plasma protein to water partition coefficient, K_p^{pl} , is solved for and substituted for the interstitial protein to water partition coefficient, K_p^{int} , assuming plasma and interstitium both consist of only protein and water [22]:

$$\frac{1}{f_{up}} = K_{\text{plasma:water}} = F_W^{pl} + F_P^{pl} K_P^{pl} \Rightarrow K_P^{int} = K_P^{pl} = \frac{1}{F_P^{pl}} \left(\frac{1}{f_{up}} - F_W^{pl} \right) \quad (7)$$

After making the substitution, the fractional volume ratio of protein in interstitium to protein in plasma, F_P^{int}/F_P^{pl} , is taken to be 0.37:

$$K_{int} = F_W^{int} + F_P^{int} K_P^{int} = F_W^{int} + \frac{F_P^{int}}{F_P^{pl}} \left(\frac{1}{f_{up}} - F_W^{pl} \right) \quad (8)$$

Acidic phospholipid and protein partitioning were calculated as [22]:

$$K_{\text{acidic phospholipid:water}} = K_{\text{neutral phospholipid}}(F_{\text{neutral}} + \alpha_+ F_+ + \alpha_- F_-) \quad (9)$$

$$K_p = 0.163 + 0.0221K_{nPL} \quad (10)$$

with cations partitioning at a rate of 20 times more ($\alpha_+ = 20$) than the neutral species into acidic phospholipid and anions partitioning 20 times less ($\alpha_- = 0.05$).

Adjustments to Membrane Affinity

The only difference in our initial method of calculating partition coefficients from the method in Schmitt [22] is in the calculation of K_{nL}^{pl} , membrane affinity. Membrane affinity data were available for 53 compounds, primarily taken from Schmitt [22], but for the remaining compounds, membrane affinity was initially predicted with a regression on neutral compounds from Endo et al. [53], reported in Wambaugh et al. [10] and used in the R package htk. This regression was roughly equivalent to using P_{ow} , the suggested approximate value in Schmitt [22]. When membrane affinity data were unavailable, they were initially predicted with a regression on logP and temperature data from Endo et al. [53] on neutral compounds, shown in (11) [10], T representing the body temperature in Celsius of the species.

$$\text{Log}(K_{\text{neutral phospholipid:water}}) = 1 - 0.0166 T + 0.882 \log P \quad (11)$$

For compounds with P_{ow} greater than 10^6 , membrane affinity values differed from P_{ow} by more than a factor of 2.

Five methods for predicting membrane affinity were evaluated before deciding to use a regression from Yun, Edginton [41], performed on the 59 compounds with membrane affinity data in Schmitt [22] (the 7 additional compounds to those evaluated only contained rabbit K_p). Alternate methods include: the original method from Poulin, Theil [23] in treating neutral phospholipids as 30% neutral lipid (K_{nL}^{pl}) and 70% water; using K_{nL}^{pl} as a surrogate for membrane affinity; and a technique similar to the calculation of neutral lipid and acidic phospholipid where the Wambaugh et al. [10] regression was multiplied by the unionized fraction and added to the ionized fraction multiplied by the Yun regression. These methods were evaluated by comparing the RMSE of the K_p without membrane affinity data, prior to adjusting protein binding. Equations for all methods evaluated are reported in Supplemental Table 2.

Adjustments to Fraction Unbound in Plasma

A lipophilic un-ionized compound should bind to the neutral lipid in plasma to a much higher degree than observed for certain compounds, based on the lipid binding model in tissue (5) and fractional volume of lipid in plasma. The error in K_p prediction for these compounds may be hypothesized to be lower due to insufficient lipid binding in the *in vitro* protein binding assay [50,54,55]. This deficiency in the assay leads to a much higher unbound concentration relative to the bound concentration for lipophilic compounds and

thus an overestimation of f_{up} and K_p . To adjust for this we recalculated the plasma:water partition coefficient ($1/f_{up}$) to account for neutral lipid binding in plasma, assuming that some lipid is absent from the *in vitro* assay. This adjustment reduces the presumably erroneous f_{up} values to what we would expect based on the lipid binding model (5) without significantly affecting f_{up} values already lower than expected or with small K_{nL}^{pl} . Many f_{up} for highly lipophilic compounds ($\log P > 5$) are reduced to below the limit of detection, which would thus otherwise be difficult to accurately measure.

In (12), the neutral lipid to water partitioning, $K_{nL}^{pl} * F_{lipid}$, is added to the plasma: water partition coefficient, $1/f_{up}$, to get an adjusted plasma: plasma water partition coefficient. This is equivalent to assuming that $f_{up}^{in vitro}$ contains the correct partitioning values except for a missing fractional volume of lipid. Although lipid is present *in vitro*, we do not know the concentration and therefore F_{lipid} is calculated as the sum of the physiological plasma neutral lipid fractional volume and 30% of the neutral phospholipid fractional volume in plasma. We use values from Peyret et al. [26] for rats and Poulin, Haddad [50] for humans. We decided for simplicity to treat 30% of the neutral phospholipid volume as neutral lipid and 70% as water rather than use our membrane affinity predictor [23,26]. We applied this correction to f_{up} and determined the improvement independent of the membrane affinity adjustment.

$$f_{up}^{corrected} = 1 / (K_{nL}^{pl} F_{lipid} + \frac{1}{f_{up}^{in vitro}}) \quad (12)$$

Adjusting K_p Prediction through Linear Regression

To further improve K_p calculation, measured K_p were regressed against the adjusted predictions for each tissue on log10-transformed K_p using R's default lm function. This allows us to substitute future K_p calculations into the regressions as K_p predictions to yield an expected measured value based on the observed discrepancies on a per tissue basis. The 95% prediction intervals were calculated as shown in Figure 4. Joint tests were performed with the estimable function from the gmodels package to calculate p-values for testing the hypothesis that the true regression is the identity line, i.e. the probabilities of the observed regressions when assuming a correctly estimated variance and a true regression with a slope of 1 and intercept of 0.

Blood: Plasma Concentration Ratio Evaluation

In vitro human B:P ratio for 279 compounds were compared to *in silico* predictions of red blood cell and B:P ratio. B:P ratio was calculated with (13) [27] from the predicted $K_{rbc:plasma}$, and $K_{rbc:plasma}$ was back calculated with the same equation from the *in vitro* B:P ratio. Both B:P ratio and $K_{rbc:plasma}$ were compared to predicted values. A regression censored below 0.1 was performed on log10-transformed $K_{rbc:plasma}$ using the function em.cens from the R package CensRegMod to incorporate negative and zero values.

$$B:P_{ratio} = F_{pt}K_{plasma:plasma} + F_{rbc}K_{rbc:plasma} = (1 - H) + H K_{rbc:plasma} \quad (13)$$

Volume of Distribution

To further evaluate the K_p adjustments and calibrations, V_{ss} predictions, calculated with initial, adjusted, and calibrated K_p values, were compared to *in vivo* human V_{ss} for 498 compounds from Obach et al. [30]. The V_{ss} were calculated with (14), where the products of K_p and their respective volumes are summed together with the plasma volume. After the initial calculation, each V_{ss} was then calculated again using the adjusted K_p values and again with the use of the regression-adjusted K_p values, excluding the red blood cell regression. Kidney and liver K_p were not adjusted to account for clearance. When the regression-adjusted K_p were used, the rest-of-body compartment was calculated as the average of the other regression-adjusted K_p . When calculating V_{ss} without regressions, rest-of-body K_p was calculated with the same method as the others, using fractional volumes for each tissue component approximated by the mean fractional volume of that tissue component across all other tissues, excluding red blood cells. Log10-transformed measured V_{ss} were regressed on predicted V_{ss} using the same method as for K_p to evaluate the model performance.

$$V_{ss} = \sum_{i \in tissue} K_i V_i + V_{plasma} \quad (14)$$

Results

Our research proceeded in three stages: 1) modification of the predictive method guided by the goal of better predictions for egregious outliers among a dataset of rat K_p , 2) empirical calibration of our modified model on a tissue-by-tissue basis, and 3) evaluation of our modified and calibrated predictions using human *in vivo* measured volumes of distribution.

Modification of the Schmitt (2008) Tissue Partitioning Model

Figure 1 demonstrates the improvement made by revising key assumptions. When the 964 tissue-specific partition coefficients (K_p), estimated from rat *in vivo* data for 143 compounds were plotted against their predicted values in Figure 1(A), outlier chemicals were identified, guiding the modifications (described below) that produced the improved predictions shown in Figure 1(B). Initially, 32 K_p for 11 compounds were over-predicted by more than a factor of 100. Including these 32, 166 K_p for 56 compounds were over-predicted by more than a factor of 10. Prediction accuracy was correlated across K_p for a given compound. Two compounds in Figure 1 (A) accounted for 11 of the most over-predicted K_p : Glycyrrhetic acid [23] had 6 measured K_p which were all predicted to be larger than 100 but were measured to be less than 0.25. The predictions for Glycyrrhetic acid improved by at least a factor of 10 due to the membrane affinity adjustment (see Methods and below). Compound JNJ17 [41] had 5 measured K_p . The initial method predicted that 2 were greater than 10,000 and an additional 3 were greater than 5,000. The protein binding adjustment improved

predictions for JNJ17 by a factor of 275. Glycyrrhetic acid and JNJ17 were two of the three compounds in our data set with logP greater than 6, which demonstrates that high logP compounds initially had the greatest over-prediction. Of the 32 K_p over-predicted by more than a factor of 100, 31 were for compounds with logP greater than 4, and 27 of the 32 K_p were for the same five compounds. The predominance of error in high logP compounds was greatly reduced after adjustments. RMSE and fold-error for all modifications are included in Supplemental Table 3.

Membrane Affinity

Membrane affinity determines the amount of chemical that partitions into lipid bilayers (e.g., cell walls) and is therefore an important parameter, albeit one for which data are often not available. We initially predicted membrane affinity using a regression based on neutral compounds [10]. We then considered four new methods: 1) using K_{nL}^{pl} as a surrogate; 2) the method of Poulin, Theil [23] that treats neutral phospholipids as 30% lipid and 70% water; 3) using a combination of our original regression for neutral compounds and the regression performed by Yun, Edginton [41] on 59 neutral, ionized, and zwitterionic compounds in Schmitt's data set; and, 4) using the Yun, Edginton [41] regression alone. Equations for all five methods are reported in Supplemental Table 2. For the 552 K_p measured for 93 compounds that lacked available measured membrane affinity, the four above methods yielded RMSE for the log10-transformed values of 0.91, 0.86, 0.85, 0.77, respectively, compared to an initial RMSE of 1.09, prior to modifying f_{up} . The K_p predicted with available membrane affinity data had a RMSE of 0.48. Among the four methods, the Yun, Edginton [41] regression performed the best because of its ionization independent minimization of phospholipid partitioning relative to P_{ow} . Schmitt [22] mentions that acidic compounds partition at a ratio of 10 to 100 times less than neutrals, but this did not account for the observed over-prediction of basic compounds.

The K_p for neutral compounds were predicted slightly but not significantly better (worsening by a mean factor of 1.1 for 55 K_p lacking membrane affinity data) with the Wambaugh et al. [10] regression than with the Yun, Edginton [41] regression. Conversely, zwitterionic compounds, which were mostly un-ionized and had a similar logP range, improved by a mean factor of 1.2. Acids and bases with the same level of neutrality (83% or more at a pH of 7.4) changed even less. All 12 neutral compounds without membrane affinity data had logP less than 3, providing little information on the validity of the adjustments for lipophilic un-ionized compounds. The 9 relatively neutral compounds (70% or more neutral) in the data set with logP greater than 3, containing 62 K_p , improved by a mean factor of 1.8. Since there was no worsening in the K_p predictions for neutral compounds, and no other methods yielded better results, we selected the regression from Yun, Edginton [41] to predict membrane affinities whenever measured data were not available.

As shown in Figure 2, the new regression for membrane affinity improved 68 K_p predictions across 18 compounds by a factor of 10, and 132 K_p across 34 compounds by a factor of 3. Of the 68 improved by more than a factor of 10, all compounds had logP greater than or equal to 3.72 and were strong bases (except 2). However, 145 K_p changed for the worse with a mean factor of 1.7. All 20 values (across 11 compounds) that worsened by more than a

factor of 3 also had higher logP values (2.88 or greater, excepting 1 K_p) and were mostly strong bases, but here there was a predominance of liver K_p and under-predictions prior to adjustment. Two compounds accounted for 10 of these values (FTY-720 and JNJ 37), suggesting possible exceptionalism or experimental error. One possible source of error for FTY-720 is its f_{up} of 0.0003, the lowest of any compound in our data set prior to modification, either in the measurement or in the assumption of only the unbound concentration penetrating the cell membrane. The 412 K_p with measured membrane affinity data were not modified, but the remaining 552 K_p had a RMSE of 0.77 after membrane affinity modification, compared to an original 1.09, on a log10 scale, improving by a mean factor of 1.9. There was consistently a similar change across tissues for each compound because of the similar fractions of neutral phospholipid across tissues.

Fraction Unbound in Plasma

For lipophilic un-ionized compounds, the actual, *in vivo* f_{up} may be hypothesized to be different than is measured *in vitro* due to insufficient lipid binding in the *in vitro* assay (i.e., less than the effective physiological quantity of lipid present *in vitro*) or other assay interference that is more pronounced for lipophilic compounds [50,54,55]. Adjusting for lipid binding in the *in vitro* assay [50,54,55] significantly reduced the estimated *in vivo* f_{up} values for those lipophilic compounds with high f_{up} that were not completely ionized, but because the majority of compounds did not belong to this class of compounds, they did not significantly change (Figure 3). The only 16 K_p (3 compounds) that changed by more than a factor of 3 improved by a very high degree. The predictions for the 5 measured K_p values for JNJ17 improved by a factor of 279, due to a logP of 7, even though the initial f_{up} was 0.02. The 10 K_p for Mazapertine (logP of 5.05 and f_{up} of 0.03) improved by a factor of 5. Amiodarone red blood cell K_p improved by a factor of 257 even though it was 92% ionized because of its logP of 7.57 (f_{up} of 0.05), demonstrating the exponential effect of the adjustment as logP increases. The adjustment method we have used is equivalent to assuming that all lipid is missing *in vitro*. In effect, some unknown amount between zero and physiologic may be missing. For those K_p with dramatic changes, the values were almost entirely determined by the amount assumed to be missing. Reducing the missing lipid by a factor of 2 yielded a change in f_{up} of nearly the same amount. Compounds that significantly changed all had high lipophilicity with a tendency toward higher f_{up} values and lower ionization. Although 336 K_p worsened, none did by more than a factor of 1.23. Conversely, 104 K_p improved by more than a factor of 1.23. RMSE changed from 0.88 to 0.79 on a log10 scale after modifying f_{up} for all K_p .

Calibration of the Model

Having modified the model for predicting tissue partitioning to reduce errors using both of the methods discussed above, we then quantified tissue-specific confidence via linear regressions of the *in vivo* data on our predicted values to estimate calibration coefficients and standard errors for K_p predictions. These estimated calibrations and standard errors establish tissue-specific levels of confidence for the prediction of chemical partitioning.

Rat *In Vivo* Tissue Partition Coefficients

Regression coefficients for all twelve tissues for which rat *in vivo* data were available are reported in Table 1, and the plotted regressions are shown in Figure 4. Additional regression coefficients using only the membrane affinity adjustment are provided in Supplemental Table 4. The regression of measured K_p on predicted K_p is significant (p -value < 0.05) in all tissues, but the correlation is weak in red blood cells, brain, and liver, all having fraction of variance explained (R^2) of 0.34 or less. Bone, lung, and spleen are predicted well enough that the regressions are not far from the identity lines. Predictions were analyzed for bias (over-or under-prediction) by testing whether the data indicated that the predictions and measurements varied from the identity line (unbiased prediction). Spleen had a p -value of 0.28, bone of 0.25, lung of 0.13, and gut of 0.02. All other tissues had p -values of 7.89×10^{-6} or less, indicating bias.

The bias in predicted K_p for adipose, brain, muscle, and skin led to predictions that mostly exceed measured K_p . The bias for kidney and liver led to predictions that are generally smaller than measured values. For liver and kidney, this is the opposite of what we expect since, considering clearance, the measured K_p would be smaller than the actual K_p given the lower concentration in tissue plasma than blood plasma. The red blood cell regression has the smallest slope, 0.27, and R^2 , 0.18, but red blood cell K_p is actually predicted more accurately than most other tissues, with 96% of values within a factor of 10 and RMSE of 0.46, compared to 0.59 for all. The regression is primarily determined by a small amount of extreme predictions relative to a limited range of measurements (mostly between 0.3 and 3). Thus we did not use the red blood cell regression in predicting V_{ss} with regressions.

Figure 5 contains a heat map of regression-adjusted K_p predicted with the revised method for all compounds with available data in humans. Example calibrations of predictions for lung, adipose, and brain for three chemicals are provided in Table 2. Tissues (rows) and chemicals (columns) were clustered hierarchically to group those that are most similar. We observe more variance between compounds than between tissues. We identify 4 overlapping clusters of compounds in Figure 5. The most distinct group is cluster 3, which consists primarily of lipophilic bases (with less lipophilic values typically having higher f_{up}). Conversely, the tenth at the left (cluster 1) consists of acids with lower f_{up} and $\log P$ (generally less than 0.2 and 4, respectively). Clusters 2 and 4 consist of a spectrum ranging from one extreme to the other with more lipophilic and neutral compounds in the cluster 4.

Significant overlap was found between clusters in Figure 5 relative to f_{up} , lipophilicity, and ionization. Lipophilicity, alkalinity, and f_{up} correlated with higher partitioning while acidity correlated with lower partitioning. These clusters are within the bounds of what we would expect given the model assumptions: With the f_{up} adjustment, an upper bound is placed on K_p for un-ionized compounds relative to $\log P$, which would otherwise partition at a much higher degree than ionized compounds, but K_p for ionized compounds lacking membrane affinity data continue to increase with $\log P$ due to the regression. However, acidic phospholipid cancels this increased effect for acids and increases it for bases, yielding the chemical clusters in Figure 5.

Figure 5 demonstrates that the regression residuals across tissues in Figure 4 should be expected to be correlated for individual compounds. Ideally, we could determine multivariate priors where K_p for each compound are consistently over-predicted or under-predicted for a sample from the regressions. The primary obstacle to using this approach is the inconsistency of available K_p data across tissues.

Human *In Vitro* Blood: Plasma Concentration Ratios

Although the rat *in vivo* data for red blood cell partitioning were limited in their ability to evaluate our predictions for that tissue, relative concentration of chemical in blood to plasma (B:P ratio) is an *in vivo* end-point that is often assayed in humans. By assuming a hematocrit fraction, K_p can be calculated from B:P ratio. Human *in vitro* B:P ratio for 279 compounds were used to further evaluate blood partitioning. Initially, inferred B:P ratio were compared to predicted values as shown in Figure 6 (A). There was no clear correlation, and the highly under-predicted red blood cell K_p values were not apparent since the B:P ratio prediction floor (0.56) is much larger for these values than the additional partitioning into red blood cells. Although hematocrit is known to vary, it was not available for the subjects corresponding to the measured $R_{b,p}$. Instead, we used a standard value of 0.44 for hematocrit in predicting B:P ratio. Thus the lowest value of K_p this allows is 0.56, but measured values fell below this value 8% of the time. Since our hematocrit value yielded negative red blood cell K_p after transformation, a censored regression was performed on log10-transformed scaled K_p , yielding a slope of 0.26, intercept of -0.22 (0.60 on plot), and standard deviation of the residuals of 0.39 (a factor of 2.5), shown in Figure 6B. Of the 279 measured red blood cell K_p , 5 were zero, and 21 were negative. 24 other K_p had prediction errors larger than a factor of 10; of these, 17 had predictions smaller, and 7 had predictions larger than inferred from measurements. There was a clear negative correlation between f_{up} and error (log10 of measured / inferred) across all K_p . The majority of the under-predicted K_p form a line roughly parallel to the identity, displaced by magnitude factor of 40. 16 of the 17 K_p under-predicted by more than a factor of 10 had f_{up} of 0.044 or less, prior to adjustment, and 8 of the 17 were acids with more than 98% ionization at a pH of 7.4. The same trend was observed for the zero and negative K_p values. Of these 26 values, 22 had measured f_{up} at or below 0.05, and 18 (17 overlapping with the 22) were more than 90% ionization at a pH of 7.4 and again acidic (except for 1 base). Ironically, strong acids were the most under-predicted as well as the most likely to be below our threshold of prediction. This fact, together with our general tendency to under-predict red blood cell K_p , suggests that we would make better predictions with a hematocrit value larger than 0.44. However, since B:P ratio is the typical model parameter of interest rather than red blood cell K_p , the prediction method does not perform as poorly as it appears. The RMSE for log10-transformed B:P ratio error is only 0.20 compared to 0.61 for red blood cell K_p with censored values excluded; 97% of B:P ratio were predicted within a factor of 3 of the measured value. Because of the weak correlation, we again did not predict red blood cell K_p with the regression in calculating V_{ss} .

Analysis of Residuals

K_p prediction error after adjusting f_{up} and membrane affinity still showed a positive correlation with neutral lipid fraction and logP for un-ionized compounds (less than 50%

ionized at a pH of 7.4). K_p over-prediction tended to increase with increasing neutral lipid fraction, and this effect is especially noticeable in adipose tissue and brain, the two tissues with the highest fractional volumes. The same trend exists with increasing logP when logP is greater than 1. This trend suggests that partitioning into neutral lipid and membrane affinity (predicted with logP for all but 50 compounds) may not be linear with logP. Alternatively, the *in vivo* measurements of K_p could be biased toward lower values for lipophilic compounds. Based on the pbtk model contained in httk, lipophilic compounds take longer to reach steady state, and 39 of the 48 compounds that were used in the K_p evaluation with available *in vitro* clearance data took longer than 1 day to reach at least 90% of the steady state concentration in humans, and this would take even longer for a model that includes diffusion-limited partitioning. Thus even studies with several hours of infusion dosing prior to measurement may not reach a true steady state for lipophilic compounds. In such cases, the K_p measured *in vivo* would be lower than its actual value.

Model Evaluation

Predicted Versus Observed V_{ss} —Having used the rat *in vivo* data to modify and calibrate our model, we turned to the human *in vivo* measured volumes of distribution (V_{ss}) reported in Obach et al. [30] as an external evaluation set. With our initial, unadjusted and uncalibrated method, V_{ss} was predicted with errors similar to the trend observed with K_p , as shown by Figure 7A. The predictions were much higher than expected from the measurements for lipophilic compounds where the larger weight of the more over-predicted tissues resulted in V_{ss} errors even greater than the initial K_p over-predictions. V_{ss} for 12 compounds were larger than measured by magnitude factor of 1000, 42 by more than 100, and 145 by more than 10.

Overall, our K_p and V_{ss} predictions greatly improved after adjusting f_{up} and membrane affinity, as shown in Figure 7B, yielding a RMSE of 0.67 compared to 1.27 previously. The two most over-predicted values improved by about a factor of 10,000,000, 11 improved by at least a factor of 100, 60 by at least 10, and 116 by 3. Predictions for 144 worsened, but only 3 by more than a factor of 3 (a factor of 6.7 at most). The four most over-predicted values, (predicted $V_{ss} > 30,000$) were primarily due to high fraction unbound relative to logP in non-ionized compounds.

V_{ss} , computed with regression-adjusted K_p (Figure 7C), showed further improvement: 113 points further improved by at least a factor of 3 (a factor of 10 at most), and 191 points worsened, 21 by more than a factor of 3. RMSE was reduced from 0.67 to 0.48. The regression line also becomes much closer to the identity, with a slope of 1.03 and intercept of -0.04 with R^2 of 0.43. The 498 compounds in the Obach data set overlapped with only 72 of the compounds used in the partition coefficient evaluation. Thus this significant improvement of volume of distribution prediction with the regressions on rat K_p provides additional validation for the calibration. Only 3 calibrated V_{ss} predictions were greater than 10 L/kg, and 95% of predictions were within an order of magnitude of the *in vivo* measurement.

Discussion

Guided by empirical evaluation against *in vivo* measured data, we have refined methods for tissue:plasma partition coefficient (K_p) predictions in order to produce more accurate estimates of K_p and volume of distribution (V_{ss}). We first analyzed a data set of 964 rat *in vivo* K_p across 143 compounds and modified our methods to address the most egregious outliers. Regressions were then performed to account for tissue-specific trends in error and to quantify uncertainty of our predictions relative to measured values. The observed frequency and magnitude of improvements in K_p prediction, in contrast to the sparsity and minuteness of worsening errors, have led us to believe that these are appropriate revisions to be adopted when using Schmitt's model. Statistical analyses of K_p predictions have been previously conducted using smaller data sets [24,22,41], however here we have taken the results of the analysis as a calibration and further evaluated these calibrated predictions. Human *in vivo* V_{ss} measurements for 498 compounds provided an independent data set allowing evaluation of our refined and calibrated method. We find that our predictions provide improved high-throughput toxicokinetic estimates.

In order for PBTK models to be applied to risk assessment, it is vital to consider the uncertainty in any estimated parameters [56,57]. Because PBTK models involve many parameters that must be inferred from experimental data, a Bayesian framework may be adopted [58–65]. In a Bayesian analysis, values of parameters known less than perfectly are characterized with “prior distributions”, probability distributions that reflect likely values for the parameters [66]. Prior distributions might be only weakly informative (e.g., partitioning is constrained between 10^{-3} and 10^3), or, at the other extreme, may be single point values (e.g., a statement like “the liver:plasma partition coefficient is 2.47”). Obtaining meaningful Bayesian estimates using weak priors can be difficult due to issues with convergence of Markov chains [66,67]. If sufficient literature data exists to construct a stronger prior, then Bayesian analysis may proceed more rapidly. In this research we describe an analysis that provides a middle ground between weak priors and infinitely precise point estimates. By empirically evaluating the performance of *in silico* predictors for partition coefficients, we have developed distributions that may be used as informative priors for Bayesian analysis [68,69] (see Table 1 and Table 2) that reflect both the information about the partition coefficients contained in the prediction model and the uncertainty (i.e., the estimated standard deviation) in those predictions reflected by failures of the models to predict real data. These distributions also may be used to assess the appropriateness of uncertainty factors when using PBTK models in any risk framework [70].

Mechanistic models represent testable hypotheses and, so far as the models reproduce reality, allow interactions with the modeled system (e.g., drug design). However, all models are by definition approximations to reality and should be expected to work in only some situations [71]. Models should be expected to have differences between observations and predictions (i.e., residual unexplained variance). Model evaluation allows for testing the assumptions and generality of the model; that is, the model is a hypothesis to be tested [72]. By quantifying and extrapolating prediction errors we are making an additional hypothesis that the drivers of these errors will also be similar in other situations. This additional hypothesis can itself be tested, as we have done using additional data for V_{ss} . We note that

although we use a factor of 10 (order of magnitude) as a point of reference for error and improvement throughout this analysis, that this is not meant to imply that something close to this amount of error is good or bad. Whether or not an amount of error is acceptable depends on the application. In our high throughput screening context, we consider order of magnitude error to be acceptable.

Calibration of K_p and V_{ss} is highly dependent on the data used in the evaluation [73]. Most of the compounds here are pharmaceuticals, potentially introducing a bias toward moieties that correspond to properties that make compounds more ideal for pharmaceutical usage, such as bioavailability. Although the adjustments we implemented improved predictions, we have not verified the actual membrane affinity of the predicted compounds to see which regression is a better predictor or the effect of adding lipid to measurements of protein binding. In particular, we have not confirmed the hypothesis that there is less than physiological lipid in the *in vitro* binding assay. It appears that for high logP chemicals the estimated f_{up} is inaccurate [50], but it is possible that other mechanisms including off-target binding (e.g., plastic, membrane) interfere with the assay for high logP chemicals and rather than correcting for missing lipid we are just estimating the partitioning into lipid, neglecting protein binding.

For these data and models, the adjustments implemented seemed the most likely and most easily corrected sources of error. Measurement and prediction error of logP and pKa also introduces another level of complexity for possible error, especially considering that these are both used in our adjustments. To maintain consistency and simplicity, we chose to predict all membrane affinities with the same regression. Although there is a difference in membrane affinity between neutral and non-neutral compounds [41], treating them separately yielded no significant differences in overall prediction accuracy. More sophisticated methods for predicting membrane affinity exist, but are not easily parameterized in a high-throughput context [74,75].

The evaluation and regressions are all based on *in vivo* measurements, but the quality of some of these measurements may be questionable [76,77]. Some *in vivo* K_p were averages of multiple measurements with differences larger than half an order of magnitude (e.g., Cefazolin, Tetracycline) [23]. If these large differences are due to experimental error, additional error is introduced into the regressions. There is additional uncertainty in applying regressions to K_p outside of the range of evaluated K_p for each tissue. Bone and skin, for example, only have evaluated K_p ranges within a factor of 300, and the red blood cell regression was not used due to a range of only a factor of 10. Higher valued K_p generally had greater logP and thus had even larger prediction errors relative to the measured K_p . The membrane affinity and f_{up} adjustments weakened this trend, and the new membrane affinity regression shifted many non-lipophilic compounds to lower values as well, maintaining a similar but improved distribution of points. This trend of over-prediction of lipophilic compounds and tissues (fat and brain) could also be due to measurement error resulting from equilibrium not being reached rather than an error in the prediction methods [78,79].

TK methods provide a critical link for interpreting chemical exposure and toxicity [6]. However, for thousands of chemicals in our environment, the necessary empirical data for

traditional approaches are lacking. The evaluated and calibrated method presented here provides predictions with quantified uncertainty, allowing confidence estimates in PBTK models for assessment of chemical risk to specific tissues.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement:

This project was supported by an appointment to the Internship/Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and EPA. The authors thank Drs. Hugh Barton and Rogelio Tornado-Velez for their helpful reviews of the manuscript. We also appreciate conversations with Drs. Weihsueh Chiu and Barbara Wetmore on protein binding.

References

1. Park YH, Lee K, Soltow QA, Strobel FH, Brigham KL, Parker RE, Wilson ME, Sutliff RL, Mansfield KG, Wachtman LM, Ziegler TR, Jones DP (2012) High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring. *Toxicology* 295 (1–3):47–55. doi:10.1016/j.tox.2012.02.007 [PubMed: 22387982]
2. Rager JE, Strynar MJ, Liang S, McMahan RL, Richard AM, Grulke CM, Wambaugh JF, Isaacs KK, Judson R, Williams AJ, Sobus JR (2016) Linking high resolution mass spectrometry data with exposure and toxicity forecasts to advance high-throughput environmental monitoring. *Environment international* 88:269–280. doi:10.1016/j.envint.2015.12.008 [PubMed: 26812473]
3. Judson R, Richard A, Dix D, Houck K, Elloumi F, Martin M, Cathey T, Transue TR, Spencer R, Wolf M (2008) ACToR—aggregated computational toxicology resource. *Toxicology and applied pharmacology* 233 (1):7–13 [PubMed: 18671997]
4. National Research Council (1983) Risk Assessment in the Federal Government: Managing the Process. In. National Academies Press, Washington (DC). doi:10.17226/317
5. Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME (2013) Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. *Toxicological Sciences* 136 (1):4–18 [PubMed: 23958734]
6. Wetmore BA, Wambaugh JF, Allen B, Ferguson SS, Sochaski MA, Setzer RW, Houck KA, Strobe CL, Cantwell K, Judson RS, LeCluyse E, Clewell HJ, Thomas RS, Andersen ME (2015) Incorporating High-Throughput Exposure Predictions With Dosimetry-Adjusted In Vitro Bioactivity to Inform Chemical Toxicity Testing. *Toxicological Sciences* 148 (1):121–136. doi:10.1093/toxsci/kfv171 [PubMed: 26251325]
7. Rovida C, Hartung T (2009) Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals—a report by the transatlantic think tank for toxicology (t (4)). *Altex* 26 (3):187–208 [PubMed: 19907906]
8. Pellegatti M (2014) The debate on animal ADME studies in drug development: an update. *Expert opinion on drug metabolism & toxicology* 10 (12):1615–1620 [PubMed: 25373428]
9. Tolonen A, Pelkonen O (2013) Analytical challenges for conducting rapid metabolism characterization for QIVIVE. *Toxicology*
10. Wambaugh JF, Wetmore BA, Pearce R, Strobe C, Goldsmith R, Sluka JP, Sedykh A, Tropsha A, Bosgra S, Shah I, Judson R, Thomas RS, Setzer RW (2015) Toxicokinetic Triage for Environmental Chemicals. *Toxicological Sciences* 147 (1):55–67. doi:10.1093/toxsci/kfv118 [PubMed: 26085347]

11. Waters NJ, Jones R, Williams G, Sohal B (2008) Validation of a rapid equilibrium dialysis approach for the measurement of plasma protein binding. *Journal of pharmaceutical sciences* 97 (10):4586–4595 [PubMed: 18300299]
12. Shibata Y, Takahashi H, Chiba M, Ishii Y (2002) Prediction of hepatic clearance and availability by cryopreserved human hepatocytes: an application of serum incubation method. *Drug Metabolism and Disposition* 30 (8):892–896 [PubMed: 12124306]
13. Jamei M, Marciniak S, Feng K, Barnett A, Tucker G, Rostami-Hodjegan A (2009) The Simcyp® population-based ADME simulator. *Expert opinion on drug metabolism & toxicology* 5 (2):211–223 [PubMed: 19199378]
14. Lukacova V, Woltosz WS, Bolger MB (2009) Prediction of modified release pharmacokinetics and pharmacodynamics from in vitro, immediate release, and intravenous data. *The AAPS journal* 11 (2):323–334 [PubMed: 19430911]
15. Wang Y-H (2010) Confidence assessment of the Simcyp time-based approach and a static mathematical model in predicting clinical drug-drug interactions for mechanism-based CYP3A inhibitors. *Drug Metabolism and Disposition* 38 (7):1094–1104. doi:10.1124/dmd.110.032177 [PubMed: 20368327]
16. O'Flaherty EJ (1981) *Toxicants and drugs: kinetics and dynamics*. John Wiley & Sons,
17. Andersen ME (1995) Development of physiologically based pharmacokinetic and physiologically based pharmacodynamic models for applications in toxicology and risk assessment. *Toxicology letters* 79 (1):35–44 [PubMed: 7570672]
18. Campbell JL, Clewell RA, Gentry PR, Andersen ME, Clewell HJ (2012) Physiologically based pharmacokinetic/toxicokinetic modeling. *Computational Toxicology: Volume I*:439–499
19. Clewell HJ, Andersen ME (1985) Risk assessment extrapolations and physiological modeling. *Toxicology and industrial health* 1 (4):111–134 [PubMed: 3843496]
20. Loizou GD, Hogg A (2011) MEGen: a physiologically based pharmacokinetic model generator. *Frontiers in pharmacology* 2:56 [PubMed: 22084631]
21. Shiran M, Proctor N, Howgate E, Rowland-Yeo K, Tucker G, Rostami-Hodjegan A (2006) Prediction of metabolic drug clearance in humans: in vitro–in vivo extrapolation vs allometric scaling. *Xenobiotica* 36 (7):567–580 [PubMed: 16864504]
22. Schmitt W (2008) General approach for the calculation of tissue to plasma partition coefficients. *Toxicology in Vitro* 22 (2):457–467 [PubMed: 17981004]
23. Poulin P, Theil FP (2000) A priori prediction of tissue: plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *Journal of pharmaceutical sciences* 89 (1):16–35 [PubMed: 10664535]
24. Payne MP, Kenny LC (2002) Comparison of models for the estimation of biological partition coefficients. *Journal of Toxicology and Environmental Health Part A* 65 (13):897–931 [PubMed: 12133236]
25. Rodgers T, Rowland M (2006) Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *Journal of pharmaceutical sciences* 95 (6):1238–1257 [PubMed: 16639716]
26. Peyret T, Poulin P, Krishnan K (2010) A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. *Toxicology and applied pharmacology* 249 (3):197–207 [PubMed: 20869379]
27. Ruark CD, Hack CE, Robinson PJ, Mahle DA, Gearhart JM (2014) Predicting passive and active tissue: plasma partition coefficients: interindividual and interspecies variability. *Journal of pharmaceutical sciences* 103 (7):2189–2198 [PubMed: 24832575]
28. Zhu X-W, Sedykh A, Zhu H, Liu S-S, Tropsha A (2013) The use of pseudo-equilibrium constant affords improved QSAR models of human plasma protein binding. *Pharmaceutical research* 30 (7): 1790–1798 [PubMed: 23568522]
29. Ingle BL, Veber BC, Nichols JW, Tornero-Velez R (2016) Informing the human plasma protein binding of environmental chemicals by machine learning in the pharmaceutical space: applicability domain and limits of predictability. *Journal of Chemical Information and Modeling* 56 (11):2243–2252 [PubMed: 27684444]

30. Obach RS, Lombardo F, Waters NJ (2008) Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 670 drug compounds. *Drug Metabolism and Disposition* 36 (7):1385–1405 [PubMed: 18426954]
31. Pearce RG, Setzer RW, Strope CL, Sipes NS, Wambaugh JF (2017) Httk: R package for high-throughput toxicokinetics. *Journal of Statistical Software* 79 (1):1–26 [PubMed: 30220889]
32. R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing,
33. Pearce R, Strope C, Setzer RW, Sipes N, Wambaugh JF (2016) httk: R Package for High-Throughput Toxicokinetics. *Journal of Statistical Software*:in press
34. R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
35. Schmitt W (2008) Corrigendum to: “General approach for the calculation of tissue to plasma partition coefficients” [Toxicology in Vitro 22 (2008) 457–467]. *Toxicology in Vitro* 22 (6):1666. doi:10.1016/j.tiv.2008.04.020
36. Rodgers T, Leahy D, Rowland M (2005) Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. *Journal of pharmaceutical sciences* 94 (6):1259–1276 [PubMed: 15858854]
37. Rodgers T, Leahy D, Rowland M (2005) Tissue distribution of basic drugs: Accounting for enantiomeric, compound and regional differences amongst β -blocking drugs in rat. *Journal of pharmaceutical sciences* 94 (6):1237–1248 [PubMed: 15858851]
38. Gueorguieva I, Nestorov IA, Murby S, Gisbert S, Collins B, Dickens K, Duffy J, Hussain Z, Rowland M (2004) Development of a whole body physiologically based model to characterise the pharmacokinetics of benzodiazepines. 1: Estimation of rat tissue-plasma partition ratios. *Journal of pharmacokinetics and pharmacodynamics* 31 (4):269–298 [PubMed: 15563004]
39. Poulin P, Schoenlein K, Theil FP (2001) Prediction of adipose tissue: plasma partition coefficients for structurally unrelated drugs. *Journal of pharmaceutical sciences* 90 (4):436–447 [PubMed: 11170034]
40. Björkman S (2002) Prediction of the volume of distribution of a drug: which tissue-plasma partition coefficients are needed? *Journal of pharmacy and pharmacology* 54 (9):1237–1245 [PubMed: 12356278]
41. Yun Y, Edginton A (2013) Correlation-based prediction of tissue-to-plasma partition coefficients using readily available input parameters. *Xenobiotica* 43 (10):839–852 [PubMed: 23418669]
42. Uchimura T, Kato M, Saito T, Kinoshita H (2010) Prediction of human blood-to-plasma drug concentration ratio. *Biopharmaceutics & drug disposition* 31 (5-6):286–297 [PubMed: 20549836]
43. Davies B, Morris T (1993) Physiological parameters in laboratory animals and humans. *Pharmaceutical research* 10 (7):1093–1095 [PubMed: 8378254]
44. Parham F, Kohn M, Matthews H, DeRosa C, Portier C (1997) Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls: I. Tissue: blood partition coefficients. *Toxicology and applied pharmacology* 144 (2):340–347 [PubMed: 9194418]
45. Paixão P, Gouveia LF, Morais JA (2012) Prediction of the human oral bioavailability by using in vitro and in silico drug related parameters in a physiologically based absorption model. *International journal of pharmaceutics* 429 (1):84–98 [PubMed: 22449410]
46. Naritomi Y, Terashita S, Kagayama A, Sugiyama Y (2003) Utility of hepatocytes in predicting drug metabolism: comparison of hepatic intrinsic clearance in rats and humans in vivo and in vitro. *Drug Metabolism and Disposition* 31 (5):580–588 [PubMed: 12695346]
47. Obach RS (1999) Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. *Drug Metabolism and Disposition* 27 (11):1350–1359 [PubMed: 10534321]
48. Strope C, Mansouri K, Clewell H, Rabinowitz JR, Stevens C, Wambaugh JF (In press) High-Throughput in-silico prediction of ionization equilibria for pharmacokinetic modeling. *Science of the Total Environment*

49. Strobe CL, Mansouri K, Kancherla J, Stevens C, Wambaugh JF High-Throughput Pharmacokinetic Modeling Using Computationally Predicted Parameter Values: Dissociation Constants. In: Society of Toxicology Annual Meeting, San Diego, CA, 2015.
50. Poulin P, Haddad S (2012) Advancing prediction of tissue distribution and volume of distribution of highly lipophilic compounds from a simplified tissue-composition-based model as a mechanistic animal alternative method. *Journal of pharmaceutical sciences* 101 (6):2250–2261 [PubMed: 22388869]
51. Snyder W, Cook M, Nasset E, Karhausen L, Howells G, Tipton I (1975) ICRP Publication 23: Report of the task group on reference man. International Commission on Radiological Protection
52. International Commission on Radiological Protection (1975) Report of the task group on reference man, vol 23 Pergamon Oxford,
53. Endo S, Escher BI, Goss K-U (2011) Capacities of membrane lipids to accumulate neutral organic chemicals. *Environmental science & technology* 45 (14):5912–5921 [PubMed: 21671592]
54. Simulations Plus (2010) GastroPlus™ Manual California, EUA.
55. Simulations Plus (2017) ADMET Predictor.
56. Barton HA, Chiu WA, Setzer RW, Andersen ME, Bailer AJ, Bois FY, DeWoskin RS, Hays S, Johanson G, Jones N (2007) Characterizing uncertainty and variability in physiologically based pharmacokinetic models: state of the science and needs for research and implementation. *Toxicological Sciences* 99 (2):395–402 [PubMed: 17483121]
57. McLanahan ED, El-Masri HA, Sweeney LM, Kopylev LY, Clewell HJ, Wambaugh JF, Schlosser PM (2012) Physiologically based pharmacokinetic model use in risk assessment--Why being published is not enough. *Toxicological sciences : an official journal of the Society of Toxicology* 126 (1):5–15. doi:10.1093/toxsci/kfr295 [PubMed: 22045031]
58. Bois FY (1999) Analysis of PBPK models for risk characterization. *Annals of the New York Academy of Sciences* 895 (1):317–337 [PubMed: 10676425]
59. Hack CE (2006) Bayesian analysis of physiologically based toxicokinetic and toxicodynamic models. *Toxicology* 221 (2):241–248 [PubMed: 16466842]
60. Tsamandouras N, Rostami-Hodjegan A, Aarons L (2015) Combining the ‘bottom up’ and ‘top down’ approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. *British journal of clinical pharmacology* 79 (1):48–55 [PubMed: 24033787]
61. Bois FY (2000) Statistical analysis of Clewell et al. PBPK model of trichloroethylene kinetics. *Environmental health perspectives* 108 (Suppl 2):307 [PubMed: 10807560]
62. Liao KH, Tan YM, Conolly RB, Borghoff SJ, Gargas ML, Andersen ME, Clewell HJ (2007) Bayesian estimation of pharmacokinetic and pharmacodynamic parameters in a mode-of-action-based cancer risk assessment for chloroform. *Risk analysis* 27 (6):1535–1551 [PubMed: 18093051]
63. Nong A, Tan Y-M, Krolski ME, Wang J, Lunchick C, Conolly RB, Clewell III HJ (2008) Bayesian calibration of a physiologically based pharmacokinetic/pharmacodynamic model of carbaryl cholinesterase inhibition. *Journal of Toxicology and Environmental Health, Part A* 71 (20):1363–1381
64. Yang Y, Xu X, Georgopoulos PG (2010) A bayesian population PBPK model for multiroute chloroform exposure. *Journal of Exposure Science and Environmental Epidemiology* 20 (4):326–341 [PubMed: 19471319]
65. Evans M, Chiu W, Okino M, Caldwell J (2009) Development of an updated PBPK model for trichloroethylene and metabolites in mice, and its application to discern the role of oxidative metabolism in TCE-induced hepatomegaly. *Toxicology and applied pharmacology* 236 (3):329–340 [PubMed: 19249323]
66. Gelman A, Bois F, Jiang J (1996) Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* 91 (436):1400–1412
67. Garcia RI, Ibrahim JG, Wambaugh JF, Kenyon EM, Setzer RW (2015) Identifiability of PBPK models with applications to dimethylarsinic acid exposure. *J Pharmacokinet Pharmacodyn* 42 (6): 591–609. doi:10.1007/s10928-015-9424-2 [PubMed: 26194069]

68. Tornero-Velez R, Davis J, Scollon E, Starr JM, Setzer RW, Goldsmith M, Chang D, Xue J, Zartarian V, DeVito MJ (2012) A pharmacokinetic model of cis-and trans-permethrin disposition in rats and humans with aggregate exposure application. *Toxicological Sciences:kfs236*
69. Tornero-Velez R, Davis J, Scollon EJ, Starr JM, Setzer RW, Goldsmith M-R, Chang DT, Xue J, Zartarian V, De Vito MJ (2012) A pharmacokinetic model of cis-and trans-permethrin disposition in rats and humans with aggregate exposure application. *Toxicological Sciences* 130 (1):33–47 [PubMed: 22859315]
70. Simon TW, Zhu Y, Dourson ML, Beck NB (2016) Bayesian methods for uncertainty factor application for derivation of reference values. *Regulatory Toxicology and Pharmacology* 80:9–24. doi:10.1016/j.yrtph.2016.05.018 [PubMed: 27211295]
71. Box GE (1979) Robustness in the strategy of scientific model building. *Robustness in statistics* 1:201–236
72. Oreskes N (1998) Evaluation (not validation) of quantitative models. *Environmental Health Perspectives* 106 (Suppl 6):1453 [PubMed: 9860904]
73. Sheridan RP, Feuston BP, Maiorov VN, Kearsley SK (2004) Similarity to molecules in the training set is a good discriminator for prediction accuracy in QSAR. *Journal of Chemical Information and Computer Sciences* 44 (6):1912–1928 [PubMed: 15554660]
74. Nagar S, Korzekwa K (2017) Drug Distribution. Part 1. Models to Predict Membrane Partitioning. *Pharmaceutical Research* 34 (3):535–543 [PubMed: 27981450]
75. Korzekwa K, Nagar S (2017) Drug Distribution Part 2. Predicting Volume of Distribution from Plasma Protein Binding and Membrane Partitioning. *Pharmaceutical Research* 34 (3):544–551 [PubMed: 27966088]
76. Khor S, Mayersohn M (1991) Potential error in the measurement of tissue to blood distribution coefficients in physiological pharmacokinetic modeling. Residual tissue blood. I. Theoretical considerations. *Drug Metabolism and Disposition* 19 (2):478–485 [PubMed: 1676658]
77. Thompson MD, Beard DA, Wu F (2012) Use of partition coefficients in flow-limited physiologically-based pharmacokinetic modeling. *Journal of Pharmacokinetics and Pharmacodynamics* 39 (4):313–327. doi:10.1007/s10928-012-9252-6 [PubMed: 22639356]
78. De Bruijn J, Busser F, Seinen W, Hermens J (1989) Determination of octanol/water partition coefficients for hydrophobic organic chemicals with the “slow-stirring” method. *Environmental Toxicology and Chemistry* 8 (6):499–512. doi:10.1002/etc.5620080607
79. Dionisio KL, Frame AM, Goldsmith M-R, Wambaugh JF, Liddell A, Cathey T, Smith D, Vail J, Ernstoff AS, Fantke P (2015) Exploring consumer exposure pathways and patterns of use for chemicals in the environment. *Toxicology Reports* 2:228–237 [PubMed: 28962356]

Figure 1:

Measured vs. predicted rat K_p comparisons before and after adjustments. (A) is of the original predictions, and (B) is after applying the adjustments to f_{up} and membrane affinity. The diagonal line is the identity. Although there is a trend, in (B), with a slope of 0.65 and intercept of 0.09 with 95% confidence intervals of (0.61, 0.70) and (0.05, 0.13) respectively, the regression of measured on predicted K_p only explains 43% of the variance. There are still many K_p for which the predictions are much larger than the measured values.

Figure 2:

Measured vs. predicted K_p after implementing a new membrane affinity regression for 552 K_p (93 compounds) without membrane affinity data. These were initially predicted with a regression on neutral compounds [53], but improved with a regression [41] on 59 membrane affinity values compiled by Schmitt (2008). The plot is colored to represent the number of orders of magnitude each point has improved ($\log_{10}$ of improvement factor). Improvement is defined as the absolute value of the initial error, $\log_{10}(\text{measured} / \text{initial prediction})$, minus the absolute value of the final error, $\log_{10}(\text{measured} / \text{final prediction})$.

Figure 3:

Measured vs. predicted K_p after adjusting f_{up} . f_{up} is reduced to take into account binding to lipid for lipophilic un-ionized compounds. The plot is colored to represent the number of orders of magnitude each point has improved ($\log_{10}$ of improvement factor). Improvement is defined as the absolute value of the initial error, $\log_{10}(\text{measured} / \text{initial prediction})$, minus the absolute value of the final error, $\log_{10}(\text{measured} / \text{final prediction})$.

Figure 4:

Regressions (solid lines) for all tissues in evaluation with 95% confidence intervals and identities (dotted lines). The regression for red blood cells was not used in the application of the regressions to the calculation of V_{ss} due to the limited range of points. Regression coefficients are included in Table 1.

Figure 5:

A histogram (left) and “heatmap” (right) showing the distribution of log₁₀ human tissue:plasma partition coefficients (K_p) calculated with the modified and calibrated method. The value of each K_p is indicated by color, with the histogram at the left indicating the value corresponding to each color and the distribution of those values. In the heatmap at the right, the columns correspond to 874 compounds while the 11 rows indicate specific tissues. The color and intensity at each location indicates the value of each K_p corresponding to the chemical and tissue indicated by column and row. Rows and columns are clustered hierarchically according to Ward’s (1963) clustering criterion, so that similar chemicals and tissues are grouped together. Clustering of chemicals is indicated by the dendrogram above the heatmap, with four specific clusters indicated by number.

Figure 6:

(A) compares human *in vitro* B:P ratio and B:P ratio calculated from predicted red blood cell K_p . (B) compares red blood cell K_p derived with Equation 13 from the same human *in vitro* B:P ratio in (B) and predicted red blood cell K_p . Both assume a hematocrit value of 0.44. Since we use this static value which does not account for biological variability and many *in vitro* B:P ratio were below 0.56, a regression censored at 0.1 was performed to predict to account for red blood cell K_p values at or below zero. The regression (solid line) yielded a slope of 0.26, intercept of -0.22 (0.60 on plot), and residual standard deviation of 0.39 on a log₁₀-transformed scale. The dotted line is the identity.

Figure 7:

Measured vs. predicted V_{ss} : A) before adjustments, B) after applying adjustments to fup and membrane affinity, and C) after applying regressions. In B and C, the plot points are colored to represent the number of orders of magnitude each point has improved (difference between initial error and final error on log10 scale). In C, the regression of measured on predicted has a slope 1.03, intercept of -0.04 , and R^2 of 0.43. The solid lines are the regressions and the dotted lines are the identities.

Table 1:

The regressions for each tissue, after f_{up} and membrane affinity adjustments, of the log₁₀-transformed measured K_p regressed on predicted K_p . The point of these regressions is to calibrate our prediction method and evaluate the prior information of these tissues.

	Intercept	Slope	Residual Standard Deviation	R-squared	Intercept Variance	Slope Variance	Covariance	N
Adipose	-0.325	0.574	0.450	0.626	0.020	0.013	-0.009	75
Bone	0.036	0.781	0.480	0.464	0.026	0.083	0.001	39
Brain	-0.167	0.543	0.647	0.249	0.015	0.022	-0.010	98
RBC	-0.006	0.267	0.286	0.182	0.014	0.057	0.002	70
Gut	0.143	0.764	0.479	0.551	0.021	0.033	-0.013	64
Heart	0.116	0.683	0.413	0.507	0.011	0.026	-0.007	105
Kidney	0.452	0.673	0.438	0.488	0.014	0.025	-0.010	103
Liver	0.475	0.621	0.563	0.337	0.015	0.025	-0.011	99
Lung	0.087	0.866	0.528	0.636	0.017	0.016	-0.010	100
Muscle	-0.022	0.658	0.373	0.568	0.011	0.023	-0.006	105
Skin	-0.090	0.566	0.279	0.573	0.020	0.050	-0.016	64
Spleen	0.034	0.765	0.574	0.402	0.025	0.066	-0.008	42

Table 2:

Example compounds with their uncalibrated and calibrated K_p . Log10-transformed regression standard deviations and chemical properties are included. See Table 1 for calibration coefficients.

	Lomefloxacin	Cyclosporin	Phencyclidine
<i>LogP</i>	0.95	2.9	4.96
<i>Ionization</i>	Zwitterion	Neutral	Base
<i>f_{up}</i>	0.72	0.083	0.47
<i>adjusted f_{up}</i>	0.72	0.074	0.26
<i>Initial Membrane Affinity</i>	1.24	823	53900
<i>Adjusted Membrane Affinity</i>	15.95	150	634
<i>Uncalibrated Lung K_p</i>	1.22	2.34	103
<i>Calibrated Lung Mean K_p</i>	1.45	2.55	67.6
<i>Calibrated Lung K_p Standard Deviation</i>	0.53	0.53	0.54
<i>Uncalibrated Adipose K_p</i>	0.45	46.8	206
<i>Calibrated Adipose K_p Mean</i>	0.3	4.31	10.1
<i>Calibrated Adipose K_p Standard Deviation</i>	0.47	0.45	0.45
<i>Uncalibrated Brain K_p</i>	0.74	5.6	31.3
<i>Calibrated Brain K_p Mean</i>	0.58	1.74	4.4
<i>Calibrated Brain K_p Standard Deviation</i>	0.65	0.65	0.65