Title Page:

Evaluation of Tissue Binding in 3 Tissues Across 5 Species, and Prediction of Volume of Distribution from Plasma Protein and Tissue Binding with an Existing Model

Frederick Hsu, Ivy Chen, Fabio Broccatelli

Affiliation: Drug Metabolism & Pharmacokinetics (FH, IC, FB), Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA.
Running Title Page:

Running title: Predicting Volume of Distribution from In Vitro Parameters

Corresponding author:

Fabio Broccatelli, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080. Tel: 6504679522, E-mail: Frederickhsu2019@u.northwestern.edu or broccatt@gene.com

Number of text pages: 34
Number of tables: 2
Number of figures: 4
Number of references: 17
Number of words in abstract: 249
Number of words in introduction: 717
Number of words in discussion: 924

Abbreviations:

AUC, area under curve; AAG, $\alpha$-acid glycoprotein; BDDCS, biopharmaceutics drug distribution classification system; CL, clearance; D, dilution factor; $F_{u\text{brain}}$, fraction unbound in brain; $F_{u\text{mic}}$, fraction unbound in microsome; $F_{u\text{p}}$, fraction unbound in plasma; $F_{u\text{tissue}}$, fraction unbound in tissue; NLS, nonlinear least square; NCA, non-compartmental analysis; PK, pharmacokinetic; $V_{d}$, volume of distribution; $V_{p}$, plasma volume; $V_{t}$, tissue volume.
Abstract:

Volume of distribution ($V_d$) is a primary pharmacokinetic parameter used to calculate the half-life and plasma concentration–time profile of drugs. Numerous models have been relatively successful in predicting $V_d$, but the model developed by Korzekwa and Nagar is of particular interest because it utilizes plasma protein binding and microsomal binding data, both of which are readily available in vitro parameters. Here, Korzekwa and Nagar’s model was validated and expanded upon using external and internal datasets. Tissue binding, plasma protein binding, $V_d$, physiochemical, and physiological datasets were procured from literature and Genentech’s internal database. First, we investigated the hypothesis that tissue binding is primarily governed by passive processes that depend on the lipid composition of the tissue type. The fraction unbound in tissues ($f_{u,\text{tissue}}$) was very similar across human, rat, and mouse. In addition, we showed that dilution factors could be generated from non-linear regression so that one $f_{u,\text{tissue}}$ value could be used to estimate another one regardless of species. More importantly, results suggested that microsomes could serve as a surrogate for tissue binding. We applied the parameters from Korzekwa and Nagar’s $V_d$ model to two distinct liver microsomal datasets and found remarkably close statistical results. Brain and lung datasets also accurately predicted $V_d$, further validating the model. $V_d$ prediction accuracy for compounds with LogD$_{7.4}$ > 1 significantly outperformed that of more hydrophilic compounds. Finally, human $V_d$ predictions from Korzekwa and Nagar’s model appear to be as accurate as rat allometry and slightly less accurate than dog and cyno allometry.
Significance Statement:

We showed that tissue binding is comparable in three tissues across five species and that the fraction unbound in tissue can be interconverted with a dilution factor. In addition, we applied internal and external datasets to the volume of distribution model developed by Korzekwa and Nagar and found comparable $V_d$ prediction accuracy between the $V_d$ model and single species allometry. Our findings could potentially accelerate the drug R&D process by reducing the amount of resources associated with \textit{in vitro} binding and animal experiments.
**Introduction:**

Volume of distribution ($V_d$) a proportionality constant between the observed concentration and the amount of drug in the body. This is used in compartmental pharmacokinetic (PK) modeling to describe the plasma concentration–time profile of drugs. While $V_d$ is an important parameter for data description, its biological relevance is not emphasized in classical compartmental modeling and PK theory. Several authors (Oie and Tozer, 1979; Rodgers and Rowland, 2007; Poulin and Theil 2009) addressed the physiological relevance of the $V_d$ term by using mechanistic modeling approaches; the aim of these studies was to describe $V_d$ in physiological relevant terms involving distribution in blood and tissues. Oie and Tozer originally described the tissue binding component by lumping binding to all tissues into a single term ($V_t$). More recent physiologically based tissue partitioning models aim to predict distribution in each major organ to better capture the shape of the PK profile. In these models different tissues are characterized based on their composition, therefore binding may vary considerably across tissues depending on the characteristics of the drug; for example in tissue partitioning models the distribution of a given compound in the adipose tissue might be predicted as substantially different from its distribute in muscles due to heterogenous characteristic of the tissues.

Lombardo *et al.* used Oie and Tozer’s model as a base to estimate $V_t$ for a large set of marketed drugs by employing simple calculated physicochemical parameters in a data driven linear model (Lombardo *et al.*, 2012). This approach delivered a fully reproducible and accurate model, which remains one of the better validated approaches in literature given the size of the dataset employed (Lombardo *et al.*, 2012). Because this model is based on calculated parameters such as pKa and lipophilicity, the effect of miscalculations for these input parameters is a considerable unknown.
Recently, Ryu et al. published a dataset of 80 compounds tested in binding experiments across different tissues and species (Ryu et al., 2020). This study highlighted the idea that binding to tissues is comparable across species and organs, in agreement with previously published work (Barr et al., 2019). Indirectly, these findings recapitulate the findings presented by Lombardo et al., which showed that while species allometry is a good predictor of $V_d$, it has limited accuracy in clearance (CL) prediction. Unlike metabolism, which is primarily determined by enzymatic processes that differ across species, distribution is primarily dominated by passive processes that depend on tissue composition and perfusion.

While the work of Ryu et al. supports tissue binding predictions from a single tissue measurement, it does not attempt to further translate these findings into $V_d$ predictions (Ryu et al., 2020). Currently, at Genentech, the only tissue binding measurements routinely available during early discovery stages is performed on microsomes. The main use of microsomal binding data is to predict the \textit{in vitro} CL of a free drug. Microsomes are artificial constructs of unsorted nature, however they maintain all the major lipid components that are believed to be relevant for tissue binding.

Recently, Korzekwa and Nagar developed a model sharing commonalities with the Oie–Tozer approach, which described distribution into tissues by using a lumped $V_t$ term, estimated based on microsomal binding (Korzekwa and Nagar, 2017); compared to the pioneering work presented by Rodgers and Rowland, this model is sensitive to changes in plasma protein binding for strong bases, and relies on a direct measurement to a biological tissue, rather than an estimate based on physico-chemical parameters (Rodgers and Rowland, 2007). This work is based on a small dataset derived from human PK experiments only. In this work, we attempt to generalize observations published by Ryu et al. and by Korzekwa and Nagar to produce a distribution model readily available during the early stages of research that can be applied.
across different species (Ryu et al., 2020; Korzekwa and Nagar, 2017). In addition, we seek to define the applicability domain of the resulting model with respect to lipophilicity, charge, and accuracy in preclinical species. Beyond increased accuracy, this methodology promises significant logistic advantages due to the reliance on a low number of in vitro measurements (microsomal binding and plasma protein binding) that are also necessary when predicting in vivo CL of metabolically eliminated compounds. These findings could support the optimization of drug half-life using in vitro (as opposed to in vivo) experiments.
Materials and Methods:

Tissue and plasma protein binding data

Datasets incorporating fraction unbound in tissue ($f_{u_{\text{tissue}}}$) measurements for brain, lung, and microsomes across three different species (human, mouse, and rat) were obtained from Genentech’s internal small molecules database; this search did not include macrocyclic compounds, therapeutic peptides, or bi-valent inhibitors. When multiple values were available, the geometric mean was adopted. Rapid equilibrium dialysis was used to determine $f_{u_{\text{tissue}}}$ as previously described (Leung et al., 2020); we performed tissue binding experiments with a 4hr incubation time, and tissue homogenates were obtained from BioIVT (https://bioivt.com/).

Calculated $f_{u_{\text{mic}}}$ values were derived using Genentech’s internal machine learning model and reported in the supplementary information (S3). Only prospective predictions (predictions run before having experimental measurements) were incorporated to avoid biasing the performance due to training set fitting. Since the model was only introduced one and a half years ago, the prospective predictions are available for 160 compounds.

All the available $f_{u_{\text{tissue}}}$ values greater than 0.001 were included in the dataset; highly bound compounds were excluded due to the experimental uncertainty typically associated with rapid equilibrium dialysis approaches (Leung et al., 2020; Chen et al., 2019).

Values for fraction unbound in plasma ($f_{u_{p}}$) greater than 0.001 and obtained in experiments for which the incubation time was 24hr were included in the dataset. Compounds that were highly bound in the same assay (>99.9%) were excluded due to the lower confidence associated with the experiment (Chen et al., 2019; Waters et al., 2008). Plasma protein binding experiments that were run with a 6hr incubation were included in additional validation sets (brain and lung binding...
datasets); due to the shorter incubation time, the adopted inclusion criteria was modified to \( f_{u_p} \) values greater than 0.1.

**Volume of distribution data**

\( V_d \) estimates from non-compartmental analysis (NCA) were performed using Phoenix® WinNonlin® version 6.4 (Certara USA, Inc., Princeton, NJ). NCA require that the plasma concentration–time profile adequately capture the area under the curve (AUC); experiments for which a substantial fraction of the AUC is extrapolated may result in less accurate quantifications of the primary PK parameters. To address this limitation, a cut-off of 20% of extrapolated AUC was applied as an inclusion criteria for experiments to be incorporated in our dataset. The estimate \( V_d \) may differ based on reference biological matrix used in the NCA analysis (blood vs plasma). This is particularly true when blood to plasma partition tends to be high for a given chemical scaffold. Historically, information about the reference biological matrix used in the NCA PK analysis has not always been made available in our corporate database. We therefore excluded scaffolds for which blood to plasma partitioning typically exceeded a value of two and for which the biological matrix used for the analysis is not known (2 projects out of 29). Only parameters derived from intravenous experiments in mouse, rat, dog, cynomolgus, and human were included in the dataset. Three datasets (brain, lung, and microsome) were used to predict \( V_d \) (S1 and S2). Brain and lung datasets included four species (cyno, dog, rat, mouse) and the microsome dataset included five species (human, cyno, dog, rat, mouse). Intravenous human \( V_d \) data was collected from Lombardo et al. (2018) or, when not available, from the DrugBank database (https://www.drugbank.ca/).

**Physiological parameters**

Plasma volume (\( V_p \)) and tissue volume (\( V_t \)) parameters for each species were obtained from literature and shown in Table 1 (Davies and Morris, 1993). Cynomolagus physiological values
were assumed to be same as Rhesus. $V_p$ is calculated by dividing the plasma volume (L) by the typical body weight of the species, while $V_t$ is calculated by subtracting total body water (L) by the blood volume (L) and dividing that difference by the typical body weight of the species. Total body water volume is a sum of intracellular and extracellular fluid and blood was not considered to be a tissue. Thus, any volume of liquid that was not blood was assumed to be tissue volume. $R_1$, as described by Korzekwa and Nagar, is the ratio of the concentration of plasma proteins in the tissue to the concentration of plasma proteins in the plasma. For neutral and acidic compounds, $R_1$ was calculated to be 0.116 in humans (60% extraplasma albumin in $V_t$ divided by 40% plasma albumin in $V_p$). Assumptions for the $R_1$ values to be used for zwitterionic species are not explicitly mentioned in Korzekwa and Nagar's paper; however, in the current work, an $R_1$ value of 0.116 was utilized under the assumption that zwitterionic compounds will predominantly bind to plasma albumin. For basic compounds, $R_1$ was calculated to be 0.052 as they are expected to predominantly bind to $\alpha$-acid glycoprotein (AAG) (40% AAG in $V_t$ divided by 60% plasma AAG in $V_p$). While $R_1$ might slightly differ from species to species, we observed that small changes in $R_1$ values have minimal impact on the results of the model. Thus, an $R_1$ value of either 0.116 or 0.052 was adopted for all species.

**Experimental and calculated physicochemical properties**

In the work published by Korzekwa and Nagar, information about the ionization class is utilized to determine the value of $R_1$. To that end, calculated $pK_a$ values were obtained using Moka (https://www.moldiscovery.com/software/moka/). Compounds were classified as basic, acidic, zwitterionic, or neutral based on the calculated charge at pH 7.4. Compounds for which the $pK_a$ value was within 0.5 units from the pH 7.4 cut-off were excluded due to the possible ambiguity in the assignment of ionic species resulting from potential errors in the calculated $pK_a$ value. Notably, the difference in $R_1$ values for different classes is relatively small; additionally the $R_1$ term becomes important for only a sub-class of compounds (highly bound with low affinity to
tissues). From a practical standpoint, assumptions on charge will most likely be important for anionic and zwitterionic compounds (typically highly bound to albumin), and unimportant for the other classes. LogD$_{7.4}$ was used as a classification cutoff and compounds without experimentally measured logD$_{7.4}$ were excluded from Genentech’s internal pre-clinical datasets. The lipophilicity assay is performed for most compounds synthesized at Genentech, therefore this further selection criteria had a minimal impact on the size of the dataset. For marketed drugs, LogD$_{7.4}$ values were collected from literature (Benet et al., 2011). When experimental LogD$_{7.4}$ was not available in the marketed drugs dataset (S1), this value was calculated using Genentech’s internal QSAR model.

**Tissue Binding Comparison and Prediction Analysis**

Using 236 unique Genentech compounds, 354 binding measurements total in either microsomes, brain, or lung tissue were compared across human, mouse, and rat. There were more fraction unbound values relative to the number of compounds because binding data was available in multiple species and tissues for certain compounds. In addition, under the assumption that different tissue matrices differ in lipid concentration but the affinity of a compound for lipids does not vary, dilution formulas (eq. 1) were utilized to estimate binding across different tissues for 352 unique Genentech compounds, yielding 399 predicted binding values. Again, there were more fu$_{tissue}$ values relative to the number of compounds due to availability of binding data in multiple tissues for certain compounds.

\[
fu_{tissue,2} = \frac{\frac{1}{D}}{\left(\frac{1}{fu_{tissue,1}} - 1\right) + \frac{1}{D}}
\]  

(1)

Dilution factors (D) were derived from non-linear regression fitting of eq. 1 using the nonlinear least squares (NLS) function (Rstudio). Once D was obtained, it was used in conjunction with fu$_{tissue,1}$ to predict fu$_{tissue,2}$.
**V_d Prediction Analysis**

For V_d prediction, Korzekwa and Nagar’s linear LK_L model (eq. 2) was used due to the simplicity of the model. The other more complicated models proposed by Korzekwa and Nagar required more inputs but did not significantly improve V_d predictions (Korzekwa and Nagar, 2017). Thus, the authors concluded that the linear LK_L model was the most appropriate model for V_d predictions. In order to allow direct comparison with the fitted parameters, fu_mic measurements at 0.5mg/mL were converted to 1mg/mL using eq. 1. Microsome, brain, and lung datasets included a total of 337, 105, and 14 compounds, respectively. For the brain and lung datasets, brain and lung fu were converted to microsomal fu utilizing the previously derived dilution factors. Finally, NLS function was used to fit eq. 2 to obtain coefficients a and b (Rstudio). V_d was then subsequently predicted with the fitted a and b values and the other parameters in eq. 2.

\[
V_d = V_p + V_t R_1 (1 - f_{up}) + V_t f_{up} + f_{up} (a \left( 1 - \frac{f_{mic}}{f_{umic}} \right) + b) \tag{2}
\]

**Statistical Analysis**

Statistical analysis included standard error for the a and b coefficients derived by non-linear fitting, R^2, AFE (eq. 3), AAFe (eq. 4), percentage within 2-fold error, and percentage within 3-fold error.

\[
AFE = 10^{average\left(\log_{10}\left(\frac{Observed}{Predicted}\right)\right)} \tag{3}
\]

\[
AAFE = 10^{average\left(\|\log_{10}\left(\frac{Observed}{Predicted}\right)\|\right)} \tag{4}
\]

R: Pearson correlations were calculated based on the log of the predicted and observed values for V_d and binding association constant K_0.
\[ K_{fu} = \frac{1 - fu}{fu} \]  

(5)

**Applicability Domain and Comparison with Allometry**

The applicability domain of the model was analyzed with respect to lipophilicity and pre-clinical allometry data. Based on the analysis presented by Benet et al., a LogD\(_{7.4}\) value of 1 can be utilized as a classification cut-off for compounds’ route of elimination; that is, compounds with a LogD\(_{7.4}\) value >1 are likely to be eliminated via the hepatic metabolic route (Benet et al., 2011). By extension, according to the biopharmaceutics drug distribution classification system (BDDCS), the distribution of compounds in this class are less likely to be affected by active transport. This is consistent with the assumptions of the distribution model introduced by Korzekwa and Nagar, which can therefore be expected to show higher V\(_d\) prediction accuracy in the high lipophilicity class. According to the same assumptions, the model can be expected to show higher accuracy in higher species (dog and cynomolgus) when a good predictivity is observed in rodents.

Lombardo et al. assessed the accuracy of allometry methodologies to predict human V\(_d\) (Lombardo et al., 2012). For compounds in our dataset for which clinical and pre-clinical data was reported by Lombardo et al., single species allometry proportionality scaling methodologies were utilized to predict human volume of distribution (Lombardo et al., 2012). This dataset was collected for the purpose of evaluating an *in vitro* only methodology to predict human V\(_d\) compared with *in vivo* methodology. Finally, the accuracy of the model with respect to the ionization class was also investigated.
Results:

**Tissue Binding**

Pfizer scientists have previously demonstrated that binding in different tissues can be extrapolated by applying simple dilution formulas (Ryu et al., 2020). While this study was rich in the number of tissues analyzed and included measurements across five species, it was limited in the size of the chemical space explored (80 unique compounds). The tissues included in the analysis (adipose, brain, heart, kidney, liver, lung, and muscle) did not include microsomal binding data, which is routinely measured in discovery phases to improve *in vitro* to *in vivo* correlations of clearance (Yang et al., 2007). The work presented by Ryu et al. highlights how tissue binding is driven mostly by non-specific binding to lipids, which are the primary components of microsomes (Ryu et al., 2020). Thus, microsomes could serve as a surrogate for binding in other tissues. In our experience, microsomes are the most frequently used biological matrix for tissues binding measurements in drug discovery, followed by homogenized brain tissues.

By extending the analysis to all the internal Genentech compounds for which binding measurements were available in either microsomes (64), brain (110), or lung (180) tissues, we were able to evaluate the variability of these 354 measurements across different species. Consistent with the findings from Pfizer and Amgen scientists, we found that tissue binding measurements are consistent across different species (Figure 1). The high correlation value ($R^2$) and low absolute average fold deviation are within the range of variability expected for experimental replicates within the same experimental conditions for a given compound. About 93% of compounds in Figure 1 have $fu_{tissue}$ within 2-fold error in the same tissue for different species. Notably, the majority of the outliers (24) are either highly bound compounds ($0.05<fu<0.01$), for which experimental determinations are less quantitative, or measurements
obtained in lung tissue, for which higher variability is typically observed due to challenges with homogenizing lung tissue (Liang et al., 2011).

Figure 2 shows the 399 predicted versus experimental $f_{tissue}$ values for 352 Genentech compounds. Non-linear fitting analysis was employed to determine the dilution factor that can be used to predict binding in a given tissue (e.g. brain) by leveraging measurements for the same compound in different tissues (e.g. microsomes). Dilution factors predicting $f_{mic}$ from $f_{brain}$, $f_{mic}$ from $f_{lung}$, and $f_{lung}$ from $f_{brain}$ were 0.0137, 0.007, and 0.59 respectively and dilution factors predicting $f_{brain}$ from $f_{mic}$, $f_{lung}$ from $f_{mic}$, and $f_{brain}$ from $f_{lung}$ were 66.6, 107.8, and 1.67 respectively (Figure 2). The analysis yielded an $R^2$ for the affinity term $K_{fu}$ of 0.76 in Figure 2a and 0.72 in Figure 2b. Interestingly, due to the asymmetrical nature of the relationship between binding affinity ($K_{fu}$) and the corresponding fraction unbound, the error in the quantitative prediction of $f_{tissue}$ observed when extrapolating from a matrix with higher lipid content (e.g. brain) to a matrix with lower lipid content (e.g. microsome) is lower compared to the opposite case. That is, 96% of $f_{tissue}$ measurements could be predicted from a different tissue with higher lipid content (dilution < 1) within 2-fold error, while 76% of measurements were within 2-fold error when $f_{tissue}$ was predicted from a different tissue with a lower lipid content (dilution >1). When the dilution value (D) is less than 1, $f_{tissue}$ predictions yield AAFE of 1.19; AAFE increases to 1.66 when the dilution value exceeds 1. These observations can be readily rationalized by looking at a theoretical example. Let us assume a measured $fu$ value of 0.4 in the diluted incubation, and a corresponding prediction of 0.8, resulting in a 2-fold deviation between the measured and the predicted $fu$. Let us now assume a value of $D=20$, the extrapolated measured $fu$ for the undiluted incubation is 0.032, while the extrapolated predicted $fu$ for the undiluted incubation is 0.17, resulting in a 5.6 fold deviation in $fu$. This can be generalized by rearranging (1) as follows:
\[ y = \frac{(1-f + \frac{1}{D}f)x}{1-fx+\frac{1}{D}fx} \]

in which \( y \) is the deviation between the predicted and the measured fu in the undiluted incubation, \( x \) is the deviation between the predicted and the measured fu in the diluted, and \( f \) is the experimentally measured fu in the diluted incubation. Taken together, these analyses highlight that a single \textit{in vitro} model can be used to fit \textit{in vivo} tissue binding from \textit{in vitro} measurements (either microsomes, lung, or brain binding), which in turn can be supplemented with plasma protein binding data to predict volume of distribution according to eq. 2, as previously proposed by Korzekwa and Nagar.

**\( V_d \) prediction with the Korzekwa and Nagar model**

To validate the linear LKL model introduced by Korzekwa and Nagar, we used the combined external (S1) and internal dataset (S2) to compare the accuracy of \( V_d \) prediction using Korzekwa and Nagar’s coefficients with the accuracy of \( V_d \) prediction using Genentech’s coefficients (Table 2). The coefficients \( a \) and \( b \) from the two methods exhibited remarkably similar values (Genentech \( a=18.22 \) and \( b=1.76 \), Korzekwa and Nagar \( a=20 \) and \( b=0.76 \)). With Genentech’s fit, 65.0% of the 337 analyzed compounds had predicted \( V_d \) values within 2-fold of observed \( V_d \) while Korzekwa and Nagar’s fit predicted 64.1% of the \( V_d \) values within 2-fold of observed \( V_d \). In addition, AAFE converged to a value of 1.9 for both sets of parameters. Given the high comparability in the statistics, the parameters originally derived by Korzekwa and Nagar were adopted to eliminate the bias resulting from evaluating and fitting a model on the same dataset.

In rows three to seven in Table 2, Korzekwa and Nagar’s fit was applied to human (n=60), cyno (n=17), dog (n=20), mouse (n=110), and rat (n=130) liver microsomal datasets. Percentage within 2-fold error ranged from 62.3% and 75.0% with rodents on the lower end of \( V_d \) prediction.
accuracy. AAFE values ranged from 1.61 to 1.94, but these were classified as accurate predictions since they all fell within 2.0.

A subset of the fu_mic dataset including 160 compounds has prospective calculated fu_mic values available (Table 2). In this dataset, the volume of distribution predictions based on experimental fu_mic (N=160, AAFE=1.96, AFE=1.25, % within 2-fold error=62.5%, % within 3-fold error=81.9%) were markedly improved compared to the predictions using calculated fu_mic (N=160, AAFE=2.21, AFE=0.88, % within 2-fold error=51.9%, % within 3-fold error=75.0%).

In addition to liver microsomal datasets, we used brain and lung datasets to further validate the hypothesis that tissue binding is comparable across different tissues and species, as well as to further validate Korzekwa and Nagar’s model. For brain (n=105) and lung (n=14) datasets, the percentage of predicted V_d values within 2-fold of observed V_d values were 69.5% and 57.1%, respectively, while the AAFE values were 1.79 and 1.84, respectively.

To assess the applicability of Korzekwa and Nagar’s model, we compared AAFE values across multiple LogD ranges and ionic species; furthermore, we utilized allometry data to assess the accuracy of the model compared to more expensive state of the art approaches (Figure 3). Accuracy in prediction observed for compounds with a LogD_7.4 >=1 (AAFE=1.80) was significantly higher compared to the accuracy observed for the more hydrophilic compounds (AAFE= 2.32). This result supports the hypothesis that lipophilic molecules primarily enter cells through passive mechanisms; less lipophilic molecules may enter cells through a variety of mechanisms including passive permeation and active transport (not captured in Korzekwa and Nagar’s model). Slight differences were observed when comparing AAFE values between different ionic species, with V_d predictions for acidic compounds being slightly less accurate.

This could also be attributed to lower lipophilicity and higher affinity for sinusoidal uptake.
transporters typically observed for acidic compounds. Overall, based on the dataset analyzed in this study, human $V_d$ predictions from Korzekwa and Nagar’s model (AAFE=1.92) appear to be as accurate as rat allometry (AAFE=1.96) and slightly less accurate than cyno (AAFE=1.71) and dog (AAFE=1.74) allometry. This result is not surprising since cyno and dog are anatomically closer to humans than are rodents.

Lastly, the accuracy of $V_d$ predictions in rodents was studied as a possible predictor of the confidence in predicting $V_d$ in higher species. When the $V_d$ prediction in rodents is within 2-fold from the experimentally observed $V_d$, the same is observed in dog or cyno in 92.5% of the cases (Figure 4). Consistently, when rodent $V_d$ predictions are not within 2-fold from the experimentally observed $V_d$, only 56.0% of $V_d$ predictions in higher pre-clinical species are within 2-fold from the experimentally observed $V_d$ (Figure 4).
Discussion:

The ability of *in vitro* and *in silico* models to predict PK properties allows us to approach the *in vivo* experiments with quantitative hypotheses. The outcome of the *in vivo* experiments may either validate these hypotheses (e.g. establish an *in vitro* to *in vivo* correlation) or identify *in vitro* to *in vivo* disconnects. These findings may increase the reliance on *in vitro* and *in vivo* models, which would reduce the need for systematic preclinical PK screening, improve the quality of chemical design, and/or point to additional experiments to characterize less understood mechanisms. Findings from early mechanistic studies to investigate disconnects in *in vitro* to *in vivo* correlations may result in the early identification of a major liability for a given chemical scaffold, allowing us to refocus chemical design with a more desirable chemical space.

Overall, quantitative hypotheses emerging from *in vitro* and *in silico* models result in saving considerable time and resources when compared to a systematic *in vivo* PK screening approach.

The importance of optimizing CL in the discovery process has been emphasized in many different publications and is well incorporated in the chemical design process in the form of *in vitro* tools, *in silico* tools, and design guidelines. Recently, the rational optimization of drug half-life has been emphasized in several publications, highlighting opportunities and unmet needs for reliable and practical *in vitro* models to be utilized in early research (Gunaydin *et al*., 2018; Broccatelli *et al*., 2019; Broccatelli *et al*., 2018). While tissue composition models significantly advanced the understanding and predictability of *in vivo* V_d (Oie and Tozer, 1979; Rodgers and Rowland, 2007; Poulin and Theil 2009), some of the key measurements that are required by these models are not readily available in the early phases of drug discovery (e.g. LogP, pKa). Furthermore, these models attempt to utilize physico-chemical properties to model binding to lipids present in tissues, rather than relying on a direct measurement of affinity to tissue components. Korzekwa and Nagar recognized that readily available f_u_mic data could be used as
a surrogate to estimate tissue binding; this approach is indirectly validated by Ryu et al., demonstrating that tissue binding is comparable across species and tissues. Our analysis based on a larger dataset of historical measurements across several tissues (microsomes, brain, lung) essentially confirm the findings of Ryu et al. We were able to derive dilution factors allowing us to convert $f_{\text{brain}}$ measurements into $f_{\text{mic}}$ estimates with high confidence (96.5% within 2-fold error) and vice versa. However, we did see a lower prediction accuracy in the former case compared to the latter case. We also observed that the experimental error propagation in the dilution formula is asymmetrical, hence the extrapolation from a matrix with lower lipid content to a matrix with higher lipid content leads to higher error. The same phenomena is to be expected when diluted plasma is used to estimate $f_{\text{u}}$ in plasma samples. Overall, these findings may contribute to decrease the resources needed to estimate binding in multiple tissue binding without appreciable information loss. The potential for a new paradigm exists in which in vitro tissue binding measurements in one species alone is enough to accurately predict tissue binding in other species and tissues.

The application of the model introduced by Korzekwa and Nagar to 456 compounds highlighted that brain or microsomal binding can be interchangeably used in conjunction with $f_{\text{u}}$ to predict $V_d$ in human and preclinical species. It is particularly encouraging that re-fitting the two model coefficients (a and b) based on the combined external (S1) and internal dataset (S2) of 337 compounds led to appreciable accuracy improvement over the original model proposed by Korzekwa and Nagar, which was based on a small set of human-only data. A closer analysis of the model accuracy stratified by ionic class and lipophilicity highlighted that the accuracy of the model for lipophilic compounds ($\log D_{7.4}>1$) approaches the accuracy of single species allometry based on dog or cyno. The accuracy of the same model for compounds with $\log D_{7.4}<1$ is considerably lower, suggesting that for these chemical entities active transporters may at times play an important role in distribution; this is in agreement with the guidelines provided by the
BDDCS system, and reinforce the expectations that the effect of drug transporters in the distribution and elimination of drugs can be expected to be important for compounds with lower lipophilicity. This simple rule of thumb may be of use when interpreting *in vitro to in vivo* correlations, and prioritizing hypothesis driven studies. Based on Genentech's internal dataset it was also possible to describe the model confidence in predicting $V_d$ for dog and monkey as a function of the accuracy for $V_d$ predictions in rodents. Not surprisingly, 92% of the $V_d$ predictions in higher species were accurate (within 2-fold) for the compounds for which $V_d$ predictions in rodents were also accurate. In the remaining cases, the accuracy of $V_d$ predictions for higher species decreased to 56%. By extension, it is reasonable to expect that good *in vitro to in vivo* correlations in rodent will translate into high accuracy in human predictions. While the model can utilize calculated $f_{u\text{mic}}$ as an input with a reasonable degree of success, predictions using experimental $f_{u\text{mic}}$ appear to be markedly better.

In conclusion, it is noteworthy to stress that the findings described in this paper provide new tools to approach human drug half-life optimization entirely based on readily available *in vitro* parameters: plasma protein binding, microsome binding, and hepatocyte stability. This could contribute to further reducing the reliance on animal experiments and accelerating the drug R&D process.
Acknowledgements:

We would like to thank Ronitte Libedinsky, Matthew Wright, Christine Bowman, Yanran Wang, and Matthew Durk for reviewing and providing editing suggestions to the paper.
Authorship Contributions:

Participated in research design: Hsu, Broccatelli

Conducted experiments: Chen

Contributed new reagents or analytical tools: Hsu, Broccatelli

Performed Data Analysis: Hsu, Broccatelli

Wrote or contributed to the writing of the manuscript: Hsu, Broccatelli
References:


Footnote
All the authors were employees at Genentech at the time the manuscript was prepared. This work received no external funding.

Legends for Figures:

**Figure 1**: Comparison of fraction unbound in tissue in three tissues across human, rat, and mouse. Species information is removed from the plot to support the hypothesis that tissue binding is comparable in a given tissue regardless of species. N=354 binding measurements. Table showing AAFE, R^2, and percentage within 2-fold error is located in the top left-hand corner of the figure. Y-axis and x-axis are presented in log scale. Solid and dotted lines represent best-fit line and 2-fold error, respectively. AAFE, absolute average fold error.

**Figure 2**: Prediction of tissue binding from a tissue with higher lipid content to a tissue with lower lipid content (a) and from a tissue with lower lipid content to a tissue with higher lipid content (b). Species information is removed from the plot to support the hypothesis that tissue binding is comparable across species and tissues. N=399 binding measurements. Table showing AAFE, R^2, and percentage within 2-fold error is located in the top left-hand corner of the figure. Y-axis and x-axis are presented in log scale. Solid and dotted lines represent best-fit line and 2-fold error, respectively. AAFE, absolute average fold error. Dilution factors predicting \( f_u_{mic} \) from \( f_u_{brain} \), \( f_u_{mic} \) from \( f_u_{lung} \), and \( f_u_{lung} \) from \( f_u_{brain} \) were 0.0137, 0.007, and 0.59 respectively. Dilution factors predicting \( f_u_{brain} \) from \( f_u_{mic} \), \( f_u_{lung} \) from \( f_u_{mic} \), and \( f_u_{brain} \) from \( f_u_{lung} \) were 66.6, 107.8, and 1.67 respectively.
Figure 3: Assessment of the applicability of the model based on $V_d$ prediction accuracy for multiple tissues, LogD ranges, ionic classes, and allometry. Number of compounds for each analysis is shown in each bar graph and AAFE values are shown above each bar graph. All species represents human, cyno, dog, mouse, and rat, while CDMR represents cyno, dog, mouse, and rat.

Figure 4: Confidence in dog or cyno $V_d$ predictions based on rodent $V_d$ predictions. Number of compounds for each analysis is shown in each bar graph and percentage values are shown above each bar graph.
**Tables:**

**Table 1.** Physiological parameters for human, cyno, dog, rat, and mouse.

<table>
<thead>
<tr>
<th>Species</th>
<th>$V_p$ (L/kg)</th>
<th>$V_t$ (L/kg)</th>
<th>$R_1$ (acid, neutral, zwitterion)</th>
<th>$R_1$ (base)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>0.043</td>
<td>0.557</td>
<td>0.116</td>
<td>0.052</td>
</tr>
<tr>
<td>Cyno</td>
<td>0.0448</td>
<td>0.6196</td>
<td>0.116</td>
<td>0.052</td>
</tr>
<tr>
<td>Dog</td>
<td>0.0515</td>
<td>0.5136</td>
<td>0.116</td>
<td>0.052</td>
</tr>
<tr>
<td>Rat</td>
<td>0.0332</td>
<td>0.614</td>
<td>0.116</td>
<td>0.052</td>
</tr>
<tr>
<td>Mouse</td>
<td>0.05</td>
<td>0.64</td>
<td>0.116</td>
<td>0.052</td>
</tr>
</tbody>
</table>

$R_1$, the ratio of the concentration of plasma proteins in the tissue to the concentration of plasma proteins in the plasma; $V_p$, plasma volume; $V_t$, tissue volume (Davies and Morris, 1993).
Table 2. Methods and statistics used to evaluate Korzekwa and Nagar’s model for predicting $V_d$.

<table>
<thead>
<tr>
<th>Method</th>
<th>Target</th>
<th>a</th>
<th>b</th>
<th>N</th>
<th>$R^2$</th>
<th>AAFE</th>
<th>AFE</th>
<th>% Within 2-fold error</th>
<th>% Within 3-fold error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNE-Liver Microsome</td>
<td>All Species</td>
<td>18.22 ± 1.39</td>
<td>1.76 ± 0.28</td>
<td>337</td>
<td>0.439</td>
<td>1.86</td>
<td>1.06</td>
<td>65.0%</td>
<td>86.0%</td>
</tr>
<tr>
<td>KN-Liver Microsome(^b)</td>
<td>All Species</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>337</td>
<td>0.446</td>
<td>1.89</td>
<td>1.16</td>
<td>64.1%</td>
<td>84.0%</td>
</tr>
<tr>
<td>KN-Liver Microsome</td>
<td>Human</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>60</td>
<td>0.700</td>
<td>1.92</td>
<td>0.98</td>
<td>65.0%</td>
<td>81.7%</td>
</tr>
<tr>
<td>KN-Liver Microsome</td>
<td>Cyno</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>17</td>
<td>0.686</td>
<td>1.61</td>
<td>0.72</td>
<td>70.6%</td>
<td>94.1%</td>
</tr>
<tr>
<td>KN-Liver Microsome</td>
<td>Dog</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>20</td>
<td>0.342</td>
<td>1.72</td>
<td>0.89</td>
<td>75.0%</td>
<td>85.0%</td>
</tr>
<tr>
<td>KN-Liver Microsome</td>
<td>Mouse</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>110</td>
<td>0.332</td>
<td>1.89</td>
<td>1.22</td>
<td>62.7%</td>
<td>82.7%</td>
</tr>
<tr>
<td>KN-Liver Microsome</td>
<td>Rat</td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>130</td>
<td>0.348</td>
<td>1.94</td>
<td>1.34</td>
<td>62.3%</td>
<td>84.6%</td>
</tr>
<tr>
<td>KN-Brain CDMR</td>
<td></td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>105</td>
<td>0.517</td>
<td>1.79</td>
<td>1.34</td>
<td>69.5%</td>
<td>85.7%</td>
</tr>
<tr>
<td>KN-Lung CDMR</td>
<td></td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>14</td>
<td>0.121</td>
<td>1.84</td>
<td>1.26</td>
<td>57.1%</td>
<td>85.7%</td>
</tr>
<tr>
<td>KN-Exp Fu(\text{mic}) CDMR</td>
<td></td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>160</td>
<td>0.291</td>
<td>1.96</td>
<td>1.25</td>
<td>62.5%</td>
<td>81.9%</td>
</tr>
<tr>
<td>KN-Calc Fu(\text{mic}) CDMR</td>
<td></td>
<td>20 ± 0.20</td>
<td>0.76 ± 0.43</td>
<td>160</td>
<td>0.113</td>
<td>2.21</td>
<td>0.88</td>
<td>51.9%</td>
<td>75.0%</td>
</tr>
</tbody>
</table>
a, first coefficient from Korzekwa and Nagar’s model; AAFE, absolute average fold error; AFE, average fold error; b, second coefficient from Korzekwa and Nagar’s model; Calc $f_{\text{mic}}$, calculated fraction unbound in microsome using physiochemical properties; CDMR, cyno, dog, mouse, and rat; Exp $f_{\text{mic}}$, experimentally measured fraction unbound in microsome; GNE, Genentech; KN, Korzekwa and Nagar; $V_d$, volume of distribution.

$^a$All species represents human, cyno, dog, mouse, and rat. $^b$KN coefficients applied to GNE data
Fraction Unbound (Tissue 1)  
Predicted Fraction Unbound (Tissue 2)  

<table>
<thead>
<tr>
<th>Category</th>
<th>AAFE</th>
<th>R²</th>
<th>% within 2-fold error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain to Lung</td>
<td>1.19</td>
<td>0.76</td>
<td>96.5%</td>
</tr>
</tbody>
</table>

Figure 2A
Figure 2B

<table>
<thead>
<tr>
<th>Category</th>
<th>Fraction Unbound (Tissue 1)</th>
<th>Predicted Fraction Unbound (Tissue 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAFE</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>% within 2-fold error</td>
<td>75.9%</td>
<td></td>
</tr>
</tbody>
</table>

- Category: Lung to Brain
- Mic to Brain
- Mic to Lung

Figure 2B
Figure 3
Figure 4

Rodent VDss Prediction is Within 2-Fold Error: 92.5%, N=40

Rodent VDss Prediction is Not Within 2-Fold Error: 56.0%, N=25