DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

DMD # 88914

AN ACCURATE IN VITRO PREDICTION OF HUMAN VD_{ss} BASED ON THE ØIE-TOZER EQUATION AND PRIMARY PHYSICOCHEMICAL DESCRIPTORS. 3. ANALYSIS AND ASSESSMENT OF PREDICTIVITY ON A LARGE DATASET.

Giuliano Berellini and Franco Lombardo*

Drug Metabolism and Bioanalysis Group, Alkermes Inc., Waltham, MA, 02451

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

RUNNING TITLE

In vitro prediction of VDss in human

* To whom correspondence should be addressed: Franco Lombardo, Alkermes, Inc., 852 Winter Street, Waltham, MA, 02451. Phone (781) 609-6514; Fax: (781) 609 5851; e-mail: franco.lombardo@alkermes.com.

Number of text pages:23
Number of tables: 9
Number of figures: 3
Number of references:32
Abstract word count: 218
Introduction word count: 750
Discussion word count: 2490

Abbreviations: FE: fold-error; fut: fraction unbound in tissues; GMFE: geometric mean-fold error; LCO: leave-class-out; $logk_{IAM}$: logarithm of capacity factor from immobilized artificial membrane columns.

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

We present a model for volume of distribution at steady-state (VD_{ss}) prediction, via fraction unbound in tissues (f_{ut}), from the Øie-Tozer equation as an extension of our and other authors previous work. It based on easily determined or computed physicochemical descriptors such as $logD_{7.4}$ and $f_{i\ (7.4)}$ (cationic fraction ionized at pH 7.4) in addition to fraction unbound in plasma (f_{up}). We had collected, as part of other work, an extensive dataset of VD_{ss} and f_{up} values and used the descriptors above, gathered from the literature, for a preliminary assessment of the robustness of the method applied to 191 different compounds belonging to different charge classes and scaffolds. After this step we addressed the use of easily computed physicochemical descriptors and experimentally derived f_{up} on the same data set and compare the results between the two approaches and against the Øie-Tozer equation using in vivo data. This approach positions itself between fully computational models and scaling methods based on in vivo animal models or in vitro K_p (tissue:plasma) data utilizing model tissues. We consider it a useful and orthogonal complement to the two very diverse approaches mentioned yet requiring minimal in vitro experimental work. It offers a relatively inexpensive, rapid, intuitive and simple way to predict VD_{ss} in human, at a relatively early stage of the drug discovery.

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Significance statement

This method allows the prediction of VD_{ss} for small molecules in human without the use of animal PK data since it utilizes only in vitro data. It is therefore amenable to use at early stages, simple, intuitive, animal-sparing and quite accurate and it may serve scaling efforts well.

Furthermore, utilizing the same dataset, we show that the performance of a model using computed pK_a and $logD_{7.4}$, still using experimental f_{up}, compares well with the model using experimentally derived values.

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

Volume of distribution, in its various forms (e.g. VD_{ss}, VD_{β}, VD_{c}), does not provide any insight into the mechanism of distribution but only a descriptive index of the propensity of a compound to partition away from the plasma compartment and it is an important determinant, together with clearance, of mean residence time (MRT) or half-life ($t_{1/2}$), the latter using VD_{β} as the volume term. There is not a "good" or a "bad" volume of distribution, and its value may range from 0.04 L/kg (plasma volume) to several hundreds of L/kg. The total body water volume is generally taken to be 0.6-0.7 L/kg and it may be considered as an upper physiological limit, thus offering a threshold value for the definition of moderate or high volume of distribution. Lombardo, Obach and Waters (2009) discussed these aspects in some detail and point out that there may be some "dominant" interactions which, in general, are governed to a large extent by physicochemical properties (Smith et al. 2015). This may explain the success in prediction upon assumption of a largely passive diffusion nature, despite hundreds or possibly thousands of specific and non-specific drug:tissue interactions. One recognized phenomenon, which may contribute to very large volumes of distribution generally observed for basic compounds (Lombardo et al 2018), is lysosomal trapping described by Daniel et al (1997) and mentioned as a possible contributor for example by Lombardo et al. (2002, 2004) and Sui et al. (2009).

The recent publication of a large dataset of human PK data (Lombardo et al. 2018) provided some impetus to revisit the prediction of volume of distribution at steady-state (VD_{ss}), using the Øie-Tozer equation (Øie and Tozer, 1979) to extract and then predict the fraction unbound in tissues (f_{ut}). The latter parameter, and the equation on which it is based, has been shown to be predictable with good results from relatively inexpensive measurements and/or computed descriptors (Lombardo et al. 2002, 2004; Sui et al 2009).

These approaches also offer access to VD_{ss} and its application to the prediction of human PK, which is discussed in more detail elsewhere (Lombardo, Obach and Waters, 2009; Lombardo et al, 2013).

There are several other methods to predict VD_{ss}. They range from scaling of animal VD_{ss} data (Obach et al., 1997; Ward and Smith, 2004; Fagerholm, 2007; Jones et al., 2011; Lombardo et al., 2013; Petersson et al., 2019), to the use of selected animal tissue as surrogate for human VD_{ss} predictions (Bjorkman, 2002) and to the use of PBPK or mechanistic approaches (Chan et al. 2018; Shimizu et al. 2019). Some authors have reported use of chromatographic indices determination from immobilized artificial membranes (Sui et al. 2009 for the prediction of f_{ut}). Other authors have coupled those chromatographic indices to binding affinity from immobilized human serum albumin on a chromatographic column for the direct prediction of VD_{ss} (Hollósy et al., 2006). In addition, the direct calculation of VD_{ss} using computed descriptors from molecular structure, without the use of any experimental parameter has been reported, among others, by Lombardo et al. (2006, 2016), Gleeson et al. (2006), Ghafourian et al. (2006), Berellini et al. (2009) and Gombar et al. (2013).

In regard to the application of f_{ut} and the descriptors used to predict it, and differing from other authors (Sui, 2009; Hollósy, 2006), we prefer the use of a well-known (and easily computed) lipophilicity parameter, such as $logD_{Oct}^{7.4}$, referred to as $logD_{7.4}$ in the rest of this work. This physicochemical parameter is much more ingrained in the use and understanding by the DMPK and Medicinal Chemistry scientists, as opposed to chromatographic indices that do not exactly reproduce $logD_{7.4}$ although they may correlate with it, and ultimately with the target property. We also utilized the experimental plasma free fraction from several experimental methods, as opposed to either calculated values, or

values obtained through chromatographic affinity determination. We believe that even though determination of f_{up} for highly bound compounds may suffer from uncertainty, it is highly preferable to logK values based on chromatographic affinity on albumin only. Along similar lines while computational approaches (Gleeson 2007) offer access to data from structure only, they do not seem to measure up to the accurate prediction level needed, especially in the case of low f_{up} (high binding). Furthermore, in more recent years, methods and tools for higher throughput f_{up} determination, such as the rapid equilibrium dialysis method (Waters et al., 2008) have become available in 96-well plates, are amenable to automation, and have become a mainstay in the pharmaceutical industry.

Methods

Human VD_{ss} and f_{up} dataset. All VD_{ss} and f_{up} data were taken from the recent trend analysis reported by Lombardo et al. (2018), referring to human intravenous data (VD_{ss}) with accompanying data for f_{up}. In all cases, the most recently reported VD_{ss} and f_{up} data were used for the modeling effort as some changes had been made, in successive publications, on some of the data reported by Lombardo et al. (2004) for the 120 cationic and neutral compounds utilized in that work.

The steps toward data collections were extensively detailed in the cited paper (Lombardo et al. 2018) as well as in previous work (Obach et al. 2008). Briefly, those data were assembled from original papers and a complete list of data, references and comments can be found in the supporting information for the respective publication, with the latter including all data from the former work. Some data were found in the literature directly as VD_{ss}, some VD_{ss} values were calculated using reported micro- or macro-constants, and some others after digitization of concentration vs. time plots via non-compartmental analysis. The plasma protein binding data reported in the cited work were taken from original references as well, and they do overall refer to multiple methods of determination, spanning across orders of magnitude. The full set of data, including $logD_{7.4}$, $f_{i\ (7.4)}$ and pK_a data, with full references, is provided as supplemental information (Supplemental Data 1- Human VD_{ss} and f_{up} dataset)

$logD_{7.4}$ and pK_a data. The experimental $logD_{7.4}$ as well pK_a data for the calculation of the fraction ionized at pH 7.4 ($f_{i\ (7.4)}$) were taken from the literature, for the initial set of 199 compounds. Overall, they were taken from different authors using different methods, but for basic and neutral compounds all $logD_{7.4}$ values were all taken from the work of Lombardo et al. (2004) and we refer to the pK_a references reported therein. In that work all

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

$logD_{7.4}$ data referred to their published $ElogD_{7.4}$ method (Lombardo, 2001) and that offers a measure of consistency. The data for the other compounds (acidic and zwitterionic) were taken from literature and they are all provided in the supporting information together with the appropriate references (Supplemental Data 1- Human VD_{ss} and f_{up} dataset). We were able to initially gather experimental $logD_{7.4}$ and pK_a data for 199 compounds, adding 79 acidic and zwitterionic compounds to the 120 basic and neutral compounds taken from literature (Lombardo et al., 2004). 8 of these compounds were excluded because of a calculated negative f_{ut} value, which cannot be transformed into a logarithmic value. We kept all other compounds in the preliminary model with 191 compounds and then also built models in turn excluding: i) the "upper outliers" (13 compounds with $f_{ut} > 1$) on a dataset of 178 remaining compounds, and ii) the 15 compounds with $f_{up} < 0.01$, on a dataset of 176 compounds. For all compounds the total anionic and cationic fractions were calculated using the sum of the contributions of each ionized species, treated independently. One quaternary ammonium compound (cephaloridine) was treated as a cation utilizing a high pK_a to ensure the generation of a highly positive $f_{i\,(7.4)}$.

Computed $logD_{7.4}$ as well as pK_a data were calculated using MoKa (v. 3.2.1, Molecular Discovery, Ltd., UK) to explore its use as in Lombardo et al. (2002) but limited to the present data set of 191 compounds to have a direct comparison with the same data.

Calculation of f_{ut} and VD_{ss} from human data. The calculation of f_{ut} was performed from human VD_{ss} and f_{up} data, using a rearranged version of the Øie-Tozer equation (Øie and Tozer, 1979) and solving for f_{ut}. The classical equation was used then to re-calculate VD_{ss} from the predicted f_{ut} values. The two equations are shown below in the order described.

$$f_{ut} = \frac{V_R f_{up}}{\left[VD_{ss} - V_P - (f_{up} V_E)\right] - \left[(1 - f_{up}) R_{E/I} V_P\right]}$$

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

$$VD_{SS} = V_P\left(1 + R_{E/I}\right) + f_{up}V_P\left(\frac{V_E}{V_P} - R_{E/I}\right) + V_R\frac{f_{up}}{f_{ut}}$$

In these equations f_{up} and f_{ut} have the usual meaning of fraction unbound in plasma and fraction unbound in tissues, respectively. The term $R_{E/I}$ refers to the ratio of extravascular to intravascular proteins, but it accounts for albumin only and it takes a value of 1.4. V_P, V_E and V_R take the values, respectively, of 0.0436, 0.151 and 0.380 L/kg, and they are defined, respectively, as the plasma volume, the extracellular fluid volume and as the physical volume in which the drug distributes minus the extracellular space (V_R, "remainder volume").

Generation and assessment of predictive performance of the models. We have utilized, as in past work from our and other authors modelling efforts, several statistics based on geometric mean fold-error on both f_{ut} and VD_{ss}, utilizing training and test sets. As in previous work by us and other authors, we utilized a rugged leave-class-out approach, and the % below 2- and 3-fold error of predicted vs. observed values. Training and test sets data are reported. All models were built using the multiple linear regression and other statistical, filter, reader and writer nodes as available in Knime (v.3.4.2, Knime, GmbH, Konstanz, Germany).

We reported, as in the past, (Lombardo et al. 2004, 2016) and as adopted by other authors (Sui et al. 2009) the performance of the leave-class-out approach (LCO) where each model is built without a class of close analogues (e.g. NSAIDs or benzodiazepines) and then the model is tested against the prediction of that class.

RESULTS

Characteristics of the Pharmacokinetic and Physicochemical Values. The data for an overall set of 199 compounds, with VD_{ss} and f_{up} available, and for which we were able to find $logD_{7.4}$ and pK_a literature values, were all taken from the work of Lombardo et al. (2018) which covers a very broad property and structural space. The description of criteria for data collection is briefly offered in Methods and, more extensively, in the work by Obach et al. (2008) and Lombardo et al. (2018). The compounds in the present data set range from a VD_{ss} of 0.04 (suprofen) to a VD_{ss} of 60 (amiodarone) L/kg, and from a f_{up} of 0.0002 (amiodarone) to a f_{up} of 0.97 (gabapentin). The heterogeneity of, and possible errors present in the data sources found is acknowledged, especially for f_{up} where different techniques have been reported in the literature, while for example all neutral and basic compounds had $logD_{7.4}$ values derived from one source, as in Lombardo et al (2004). Structural-therapeutic classes were also identified for further analysis as reported in previous work, and no class was considered unless it comprised at least 10 analogues.

Model building. Models 1 and 2 (Model 1 shown in Equation 1 below) were generated using the available experimental data on 199 compounds, one including all but 8 compounds with $f_{ut} < 0$ (dataset of 191) and the other excluding, in addition, 13 compounds with $f_{ut} > 1$ (dataset of 178), respectively. They were built as preliminary models based on experimental $logD_{7.4}$ as well as f_{up} and cationic $f_{i\,(7.4)}$ from experimental pK_a, to assess their predictive performance using several statistics reported in Tables 1 and 2.

$$\boldsymbol{logf_{ut} = -0.249 * logD_{7.4} \quad -0.999 * f_{i\,(7.4)} + 0.735 * logf_{up} + 0.070} \textbf{ (Eq. 1)}$$

The sign of the parameters, as observed in previous work (Lombardo et al. 2004; Sui et al. 2009) is intuitively what should be expected since a very high fraction ionized (for a base) should be detrimental to free fraction in tissues, the electrostatic interactions with membranes phospholipids being a very significant determinant of a compound behaviour. Lipophilicity should show the same sign although the value of the coefficient, not being scaled and ranging at least an order of magnitude (some 13 $logD_{7.4}$ units), shows a lower value. Plasma protein binding, conversely, would limit access to tissues and membranes (although there is albumin in addition to many other proteins in tissues) and its coefficient is indeed positive. Table 1 and 2 show the coefficients and relevant statistics confirming the extremely high relevance of all three parameters. They also show that there was not much difference whether the models were built with or without the inclusion of f_{ut} values above 1. Table 3 shows the performance of the models in the prediction of f_{ut} and back-calculation of VD_{ss} from predicted values, and it reports the statistics on Model 3, which was built using the 176 compounds with f_{up} values at or above 0.01.

We did explore the use of fraction ionized for anionic groups (whether anions or zwitterions, data not shown) as a separate term but we did not find it to be significant in the initial models, with experimental lipophilicity and pK_a data. These results suggest that the anionic charge fraction is not a needed descriptor, at least for our data set, and we did not pursue its application any further. Sui et al. (2009), on the other hand, included both charges (only the cationic charge for zwitterions) in the single charge descriptor they used.

The coefficient the reported by Sui et al. (2009) for the $logk_{IAM}$ index is closer to our $logD_{7.4}$ coefficient (-0.3199 vs. -0.249, respectively, Model 1), than either of the other two other coefficients, which were reported to be smaller (taken as absolute values) than ours, with 0.4699 vs. 0.735 for $logf_{up}$ and -0.4069 vs. -0.999 for $f_{i\ (7.4)}$, respectively. The

intercept (error) is not significant in our Model 1 ($P > |t| = 0.333$), while it was significant in the model we explored using both charge types as independent descriptors.

As a test, we calculated the f_{ut} and VD_{ss} for the 8 compounds we excluded because of a negative f_{ut} value, using Model 1. The statistics are in Table 4 and we note here that they are all anionic compounds with very small VD_{ss} values (in some cases confined to blood or plasma, with VD_{ss} of 0.08 and 0.04 L/kg, respectively) and the overall GMFE on the test set is 3.04. This is due to significant outliers (e.g. glyburide with a VD_{ss} FE value of 6.26) which weigh heavily in a small test set. The entire set of compounds with observed and predicted values is reported (Supplemental Table 1). The overall outcome of this test, however, was unsatisfactory.

As a second step we calculated the predicted f_{ut} and VD_{ss} values for the 13 compounds with $f_{ut} > 1$ utilizing Model 2, which was built with their exclusion. The full results are shown in the supplemental information (Supplemental Table 2) and the GMFE for VD_{ss} on the test set was 1.95 while the bias (observed-predicted) was found to be -0.23, as reported in Table 4. We note that all compounds in this set have experimental VD_{ss} values < 0.5 and that the prediction does a reasonably good job in keeping the GMFE of VD_{ss} prediction just below 2-fold but with a much larger GMFE for f_{ut} at 5.6. The two largest values of fold-error were found for cephradine to be 3.64 and enalaprilat 3.13 (Supplemental Table 2).

We attempted to remove all compounds with a $f_{up} < 0.01$, based on recent guidance from FDA for in vitro DDI studies (FDA guidance document, October 2017) out of concern about the accuracy of such measurements. Also, we considered the simulations reported by Waters and Lombardo (2010) regarding the sensitivity of f_{ut} on f_{up} and $R_{E/I}$, especially when looking at compounds with $f_{up} < 0.1$. Lombardo et al. (2002) had also explored, on a small

test set of 14 proprietary compounds, the exclusion of compounds with $f_{up} < 0.02$, and reported significantly improved mean fold-error values on the prediction, although their usable set was reduced, in some instances, to 6 compounds, when compounds with such low f_{up} were excluded. When we performed a similar test, excluding the very highly bound compounds from the 191 compounds dataset used for Model 1, and recast the model, now termed Model 3, the latter yielded a reasonably good result. The GMFE for f_{ut} and VD_{ss} on the training set for the 176 compounds model were 2.10 and 1.73, respectively, as shown in Table 3. These values are almost identical to the values for Model 1. The test set (Supplemental Table 3), represented by the 15 compounds with $f_{up} < 0.01$ yielded a GMFE of 2.20 for VD_{ss} as shown in Table 4 which is a bit higher than the GMFE for the prediction of compounds with $f_{ut} > 1$ (Table 4).

We also performed what we consider a very rugged test, the leave-class-out (LCO), which we and other authors have utilized in several examples of predictive work (Lombardo et al. 2004, 2016; Sui et al. 2009). In this approach, all members of a class of analogues (at least 10 for each class) are removed, and the model is built without them. Then each of model is used to predict the class of analogues not used in deriving it. The results are shown in Table 5 and the overall GMFE was a very good 1.69 with 68 and 89 % of compounds predicted with 2- and 3-fold, respectively.

In addition, we performed a test utilizing 22 of the 60 compounds which overlapped with the set used by Lombardo et al. (2013) in their scaling work utilizing the Øie-Tozer method based on all three species, as in model V7 in that work. Two compounds were then excluded in keeping with their approach of using only compounds with in vivo $0 < f_{ut} < 1$ and we recalculated the GMFE for those 20 compounds, from the available supporting information using model V7, obtaining a value of 1.44. The GMFE calculated from Model

1 predicted VD_{ss} (training set) yielded a value of 1.36. The full set of data is reported in the supplemental information (Supplemental Table 4). Most recently, Petersson et al. (2019) have revisited and discussed the use of f_{ut} (from the average of three species) as a predictor of VD_{ss} in human with and without the elimination of aberrant f_{ut} values. In analogy with our conclusions they recommended it as the most accurate method, at least at later stages, when data in rat, dog and monkey are available.

Armed with these results, generated using experimentally determined $logD_{7.4}$, pK_a and f_{up} values, we set out to explore the use of computed $logD_{7.4}$ and pK_a values, using MoKa, for the same 191 compounds we used to develop Model 1 (Equation 1). Model 1c was built and its statistics (on the 191 compounds of the training set) are reported in Table 6. We note that the coefficients of the equation are very similar to the ones in Model 1 (Equation 1) and that the observed GMFE values for f_{ut} (2.36) and VD_{ss} (1.86) for the same training set of 191 compounds are only slightly higher than the values reported in Table 3 for Model 1. In addition, the model shows a greatly increased accuracy with respect with the data reported by Lombardo et al. (2002) which were based on significantly smaller dataset (64 compounds) comprising only basic and neutral compounds. Both outcomes, however, were obtained after recalculation of training set values, and all 64 compounds were comprised within the 191 compounds set.

A LCO approach utilizing computed $logD_{7.4}$ and pK_a values was also tested and the results are in Table 7. Overall, the performance is like the one observed for Model 1 (Table 5) even though there are some noticeable differences between models. For example, β-lactams perform better with the former (all in vitro data) and benzodiazepine perform better with the latter model (computed $logD_{7.4}$ and pK_a). Similarly, the use of computed descriptors (Supplemental Table 4) did not seem to worsen the performance and the overall

GMFE for the 20 compounds mentioned earlier (used for the in vivo methods comparison) is 1.44, which is essentially the same of the value obtained from Model 1 and identical to the recalculated value using in vivo literature data (Lombardo et al. 2013, model V7 in that work). It is recognized, however, that the test set is relatively small.

We also examined the performance of Model 1 across the ranges of predicted f_{ut} values and the 4 charge classes and the results for the latter are shown in Table 8. We note that the performance (recalculated values from Model 1) is not highly variable by the charge class and indeed anions, the class with generally low VD_{ss}, and zwitterions are predicted very well. Thus the "homogeneity" of prediction is generally preserved across charge classes. These observations are generally confirmed graphically by the plots in Figures 1 (compounds shaded by f_{ut} ranges) and 2. In Figure 1 we show the observed vs. predicted VD_{ss} value and we note that there is some variation (generally underprediction) at higher rather than lower VD_{ss} (and predicted f_{ut}) values. In Figure 2 we show the same compounds colored by their charge class and the red dashed vertical line is set at 0.7 L/kg or total body water, on the X-axis (predicted VD_{ss} values). There are 65 compounds with predicted VD_{ss} < 0.7 L/kg and the GMFE is 1.59 with 75 and 91 % of compounds below 2- and 3-fold error, respectively. The blue line, instead, identifies an (arbitrary) upper limit of 2.8 L/kg approximately equal to 4 times the total body water, with 52 compounds above that threshold. In this range the GMFE is 1.96, and the corresponding fold value thresholds are 58 and 81%.

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

DISCUSSION

We start our discussion from the exploration of the anionic fraction and we note that Sui et al. (2009) did not report the use of a specific anionic fraction term. They used only one $f_{i\,(7.4)}$ term in their equations, treating the zwitterionic compounds (6 in the training and 1 in the test set) as cations (with a positive sign of the values), and the anions as such with a negative sign for the latter $f_{i\,(7.4)}$ values. They also used a chromatographic index and a smaller data set (121 compounds), with a somewhat lower range of VD_{ss}, which may have influenced the significance of the charge state, and the overall magnitude of coefficients. Our coefficients for $logf_{up}$ and $f_{i\,(7.4)}$ are in fact significantly different from theirs (see Results). We did try, as mentioned in the Results section, the incorporation of a separate term for anionic charge fraction but we did not find it necessary. In addition, the coefficients of Model 1 are very close to the coefficients reported by Lombardo et al. (2004) for 120 neutral and basic compounds only (set entirely contained within the 191 compounds used). In that work, the authors reported values of -0.2294 ($ElogD_{7.4}$), 0.8885 for $logf_{up}$ and -0.9311 for $f_{i\,(7.4)}$. This observation suggests that the fraction ionized for anionic groups may not be strongly correlated with f_{ut}, even after the inclusion of a sizable number of anionic and zwitterionic molecules. It is possible that the f_{up} and $logD_{7.4}$ terms for anionic compounds, considering their higher propensity toward protein binding (largely but not exclusively to albumin), may be able to explain the smaller variance in f_{ut} (and VD_{ss}) for these compounds. At any rate, as we did not find the anionic $f_{i\,(7.4)}$ to be necessary and, at least within the domain of physicochemical properties, range of VD_{ss} values, and structural features expressed by our dataset, there would be no need to determine it experimentally for acidic compounds.

As described in Results we tried to predict the VD_{ss} of the 8 compounds with a negative f_{ut} value which we had set aside from the overall set of 191 compounds. The results, shown in Table 4 (statistics in Supplemental Table 1) for the full set yielded a relatively poor performance with only 1 compound (naproxen) predicted at a FE of 2 and all other above, for an overall GMFE of 3.04. This set is of course a very harsh test, as it may be expected, and the model cannot effectively compensate for the negative values obtained through the rearranged Øie-Tozer equation, its basic assumption being passive diffusion. A poor performance, in our experience, is sometimes observed when data from animal studies with a back-calculated $f_{ut} < 0$ and $f_{ut} > 1$ are used as species to species differences seem to matter significantly. That is the basis of the selection of the 38 compounds by Lombardo et al. (2013) all having $0 < f_{ut} < 1$. The prediction using the present model(s) model will always generate positive f_{ut} values which will offer no potential warning as it may be the case with methods using animal data. Conversely, very recent results such as those reported by Petersson et al. (2019) seem to indicate that even with $f_{ut} < 0$ and $f_{ut} > 1$ values a good overall prediction can be generated. Single compounds may have to be examined though, via the generation of more data at later stages. This will involve the use of animal data and much more detailed studies (e.g. transporters) which is much more expensive and involved, and it is reserved for late(r) stage candidates.

We then turned our attention to the calculation of the f_{ut} values for the compounds having $f_{ut} > 1$. Such values, f_{ut} being a "fraction" are also considered an aberrant product of the rearranged Øie-Tozer equation. It may be reasonable to expect in general a predicted f_{ut} (much) smaller than a value calculated from the Øie-Tozer equation, being that these compounds were excluded a priori. Nevertheless, we obtained some predicted f_{ut} values > 1 and, in most cases, they yielded reasonably close and acceptable predictions of VD_{ss}.

This may lend support to the fact that $f_{ut} > 1$ values may still generate good VD_{ss} predictions. That is, a seemingly aberrant predicted f_{ut}, which may caution against its use, may in fact be usable in the prediction. We caution, though, that it is difficult to judge the validity of such results, in the absence of corroborating supplementary data. At any rate, the predictive GMFE from Table 4 was 1.95, while the results for the full dataset are also reported (Supplemental Table 2).

The argument based on transporters, as a possible explanation for either type of aberrant results, i.e. $f_{ut} < 0$ and $f_{ut} > 1$, offered by Waters and Lombardo (2010), may be of difficult application for prospective predictions. This may be the case even if transporters data and/or observation from in vivo PK in animals were available for the compounds being examined. Grover and Benet (2009) showed that transporters could be important, especially at the organ level and primarily in liver and kidney, and they can influence VDss, but their effect is generally limited to 2-fold and varies greatly from species to species. Furthermore, the impact of transporters, as it may be intuitively understood, is different depending on whether they are efflux or uptake ones and depending on the type of volume of distribution considered. Smith et al. (2015) more recently reiterated the fact that transporters do not seem to be major determinants of volume of distribution, even though there are notable exceptions. The latter authors note that charge (first and foremost) and then lipophilicity are the primary determinants of volume of distribution.

Thus, it may be more likely that the empirical nature of the Øie-Tozer equation, coupled with the choice of fixed ($R_{E/I}$) or species-dependent terms, plus the uncertainties associated with the determination of f_{up} especially when very low, are the causes of $1 < f_{ut}$ and $f_{ut} < 0$ values. Lombardo and Waters showed the impact of the $R_{E/I}$ term, by simulating

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

f_{ut} back-calculation response when varying its value, for compounds with relatively high and relatively low f_{up}. In the context of the present work, the test set is too small to allow a definitive decision as only a few (4/13) predicted $f_{ut} > 1$ values were observed (Supplemental Table 2). We add that we generally have found the use of animal $f_{ut} < 0$ and $f_{ut} > 1$ values detrimental to a good performance toward predicting human VD_{ss}.

As a third approach we examined the prediction of compounds with $f_{up} < 0.01$ in part based on the findings of Waters and Lombardo (2010) on the sensitivity toward $R_{E/I}$ of back-calculated f_{ut} for those compounds, and in part based on FDA guidelines for in vitro DDI studies (FDA Guidance Document, October 2017). Model 3 did show similar performance on recalculated values for the training set as Model 1 did. Its predictive GMFE based on the excluded (test) compounds only, yielded a value of 2.20 while, for example, the prediction of compounds with $f_{ut} > 1$ yielded a GMFE of 1.95 (Table 4, full set in Supplemental Table 3). Also, the prediction yielded a very respectable % < 2-fold value of 73% but identical to the % < 3-fold. That is, no compound was predicted between 2- and 3-fold and outside the narrower limit a larger error was observed. This may suggest caution in predicting VD_{ss} values for compounds with f_{up} values (known a priori as required by the model) below 0.01.

The next step was the prediction of the compound classes using the LCO approach as described in Results. In general, the results show a very good performance across many classes with the tricyclic antidepressant being the only class with GMFE > 2 which may be due to the difficulty, in general, to predict very high volumes of distribution. Excellent results were obtained for steroids, adrenergics and fluoroquinolones. The NSAIDs, which have generally relatively low volumes are overall well-predicted (GMFE 1.88) but with an inferior performance for the 2-fold range and a respectable 83% within 3-fold. We also

tested the model against some of the in vivo data reported by Lombardo et al. (2013) utilizing the 20 overlapping compounds with $0 < f_{ut} < 1$. The animals are naïve to each of the compounds administered, but the statistical and scaling methods are training set dependent. Thus, the approach is a fair comparison with other scaling methods. We obtained a GMFE of 1.36 vs. a GMFE of 1.44 for the in vivo prediction using three animal species. We point out, at any rate, that cost and ethical considerations, as well as time and amount of available material, weigh heavily in these comparisons and they are clearly in favour of in silico and in vitro methods. Furthermore, even outside the use of computed descriptors, we note that when scaling of human PK prediction is needed (generally at later discovery stages) all experimental data needed should have been long generated for those and even earlier analogues. And that this approach positions itself as an orthogonal and inexpensive one between in silico methods (based on structure only) and methods such as Øie-Tozer and PBPK utilizing extended in vivo data.

Along the same lines discussed above and illustrated in the Results section, we performed a LCO test using computed $logD_{7.4}$ and $f_{i\ (7.4)}$ and systematic removal of all analogues of a class to be predicted. The results (Table 7) are comparable to the results obtained with the model based on experimental values (Table 5) and while there is some decrease in accuracy for β-lactams and tricyclic antidepressant, there is an improvement on benzodiazepines. Furthermore, we have performed a test using the same 20 compounds described above with in vivo data in three species and we found, albeit within the limit of the small set, and considering that these are recalculated from Model 1c, that its performance (GMFE 1.44) in on par with the in vivo (GMFE 1.44) and the in vitro Model 1 (GMFE 1.36). The data for the full set are available in Table S4.

Clearly, the accuracy of the computational method used for logD7.4 and pKa calculation is of paramount importance as the performance may vary, from class to class and across a range of structures, which the model may or may not have been well-parametrized for. It is advisable to test the prediction with a few "probe" compounds with same scaffold of the compounds of interest. Several available computational methods, whether commercial or in-house, are amenable to training with proprietary compounds data, and that is an improvement that should be taken advantage of, in the prediction of pK_a and $logD_{7.4}$ for any application. At any rate good quality f_{up} values will still be needed, if the Øie-Tozer approach is used.

Lastly, we discuss observations on the range of applicability of the method toward the prediction of VD_{ss} as shown by Figure 1 and Figure 2 in addition to Table 8. Figure 1 shows the observed vs. predicted VD_{ss} correlation across the entire range, and it is apparent that there are underpredicted values as the (predicted) VD_{ss} increases. This may caution toward its application at very low f_{ut} values and generally very high VD_{ss} values and may require, in future developments, more and/or quadratic terms. The f_{ut} value ranges are identified by the shadings. Generally, compounds with very large VD_{ss} like tebufelone (12 L/kg), maprotiline (45 L/kg) and amiodarone (60 L/kg) are significantly underpredicted as exemplified by some of the reported tabular data. Figure 1 and Table 8, on the other hand, shows the performance of Model 1 when compounds are segregated (but not excluded in casting the model) according to charge class. We note that anions, generally having lower VD_{ss} values, as well as zwitterions are predicted quite well, with a low maximum fold-error value and very high number of compounds with 2- and 3-fold, thus reinforcing the observed high number of good prediction in the lower range of the plot in Figure 1. The

same data (charge classes) is presented in Figure 2 where compounds are now colored according to their charge class.

To define the domain of extrapolative failures when using PK scaling methods, Jolivette and Ward (2005) took the approach of calculating descriptors such as hydrogen-bond donors and acceptors number as well as logP. They divided VD_{ss} values in three bins, using 0.7 and 3.5 L/kg as thresholds, based on the animal data, and differentiated the results among rat, dog and monkey. Similarly, we looked at the predicted value for the 191 compounds set by Model 1 and identified similar thresholds (0.7 and 2.8 L/kg), to identify bias and accuracy of prediction based on range of values. In doing so we overlaid two vertical lines onto the initial plot of Figure 2 (colored by charge class) using those thresholds. The numerical data are shown in Table 9. In general, the performance decreased by all indicators (GMFE, % within fold-error and bias) as VD_{ss} increased, but it remained reasonably good. This is a different way to show that indeed larger volumes are more difficult to predict (as in Figure 1), and anions, generally residing within low(er) volume ranges, seem more easily predicted than cations. Lastly, we generated a 3D plot using $logD_{7.4}$, f_{up} and FE (compound colored by the latter quantity) to identify the numerical values that might yield less accurate prediction. This is presented in Figure 3. We show that combination of very low f_{up} and high $logD_{7.4}$ tends to increase FE. That is, a highly bound and lipophilic compound, will likely not be accurate, while each of the two experimental values may not, by itself, necessarily yield a high FE.

The main aim of this work was to explore the predictive power and the limitations of an in vitro method to predict VD_{ss}, which would yield a good performance with easily determined experimental parameters, and which could position itself between much costlier and resource demanding in vivo and fully in silico (structure only) methods. We

note that for many analogues, let alone compounds approaching clinical candidate status, $logD_{7.4}$, pK_a and f_{up} data should be available and thus this prediction does not require variables other than those that should be routinely measured. We also note that, while f_{up} is not a parameter that should be optimized, its determination is amenable to 96-well plate and it is relatively routinely performed to explain PK-PD relationship as well as for application such us unbound concentration ratios in brain and plasma. We find this approach, using easily determined physicochemical descriptors (or computed with a control on accuracy), to be quite accurate and generally on par with other methods, orthogonal to several of them, easy to utilize, relatively inexpensive, intuitive and, very importantly, animal sparing and we believe it should find application in human PK prediction at early stages of Discovery.

Author Contributions

Participated in research design: G. Berellini and F. Lombardo

Conducted experiments: N/A.

Contributed new reagents or analytic tools: N/A.

Performed data analysis: G. Berellini and F. Lombardo.

Wrote or contributed to the writing of the manuscript: G. Berellini and F. Lombardo

References

Berellini G, Springer C, Waters NJ, Lombardo F (2009) In silico prediction of volume of distribution in human using linear and nonlinear models on a 669 compounds data set. *J Med Chem* **52:** 4488–4495.

Bjorkman S. (2002) Prediction of the volume of distribution of a drug: which tissue–plasma partition coefficients are needed? *J Pharm Pharmacol* **54**:1237-1245.

Chan R, De Bruyn T, Wright M, Broccatelli F (2018) Comparing Mechanistic and Preclinical Predictions of Volume of Distribution on a Large Set of Drugs *Pharm Res* **35**: 87.

Daniel WA, Wojcikowski J (1997) Contribution of lysosomal trapping to the total tissue uptake of psychotropic drugs. *Pharmacol Toxicol* **80:**62– 68.

Fagerholm U (2007) Prediction of human pharmacokinetics evaluation of methods for prediction of volume of distribution. *J Pharm Pharmacol* **59:**1181–1190.

FDA regulatory guidance (2017) https://www.fda.gov/regulatory-information/search-fda-guidance-documents/vitro-metabolism-and-transporter-mediated-drug-drug-interaction-studies-guidance-industry. Accessed May 24, 2019.

Ghafourian T, Barzegar-Jalali M, Dastmalchi S, Khavari-Khorasani T, Hakimiha N, Nokhodchi A (2006) QSPR models for the prediction of apparent volume of distribution. *Int J Pharm* **319:**82–97.

Gleeson MP, Waters NJ, Paine SW, Davis AM (2006) In silico human and rat Vss quantitative structure-activity relationship models. *J Med Chem* **49:**1953–1963.

Gleeson MP (2007) Plasma protein binding affinity and its relationship to molecular structure: an in-silico analysis. *J Med Chem* **50:**101–112.

Gombar VK, Hall SD (2013) Quantitative structure-activity relationship models of clinical pharmacokinetics: clearance and volume of distribution. *J Chem Inf Model* **50:**948–957.

Grover A, Benet LZ (2009) Effects of Drug Transporters on Volume of Distribution. *AAPSJ*, **11:**250-261.

Hollósy F, Valko K, Hersey A, Nunhuck S, Keri G, Bevan C (2006) Estimation of Volume of Distribution in Humans from High Throughput HPLC-Based Measurements of Human Serum Albumin Binding and Immobilized Artificial Membrane Partitioning. *J Med Chem* **49:** 6958-6971.

Jolivette LJ, Ward KW (2005) Extrapolation of human pharmacokinetic parameters from rat, dog, and monkey data: molecular properties associated with extrapolative success or failure. *J Pharm Sci* **94:**1467–1483.

Jones RD, Jones HM, Rowland M, Gibson CR, Yates JW, Chien JY, Ring BJ, Adkison KK, Ku MS, He H, Vuppugalla R, Marathe P, Fischer V, Dutta S, Sinha VK, Bjoernsson T, Lave T, Poulin P (2011) PhRMA CPCDC initiative on predictive models of human pharmacokinetics, Part 2: Comparative assessment of prediction methods of human volume of distribution. *J Pharm Sci* **100:**4074-4089.

Lombardo F, Shalaeva MY, Tupper KA, Gao F (2001) ElogDoct: A Tool for Lipophilicity Determination in Drug Discovery. 2. Basic and Neutral Compounds. *J Med Chem* **44:**2490–2497.

Lombardo F, Obach RS, Shalaeva MY, Gao F (2002) Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data. *J Med Chem* **45:**2867–2876.

Lombardo F, Obach RS, Shalaeva MY, Gao F (2004) Prediction of human volume of distribution values for neutral and basic drugs. 2. Extended data set and leave-class-out statistics. *J Med Chem* **47:**1242–1250.

Lombardo F, Obach RS, DiCapua FM, Bakken G, Lu J, Potter DM, Gao F, Miller MD, Zhang Y (2006) A hybrid mixture discriminant analysis-random forest computational model for the prediction of volume of distribution in human. *J Med Chem* **49:**2262–2267.

Lombardo F, Obach RS, Waters NJ (2009) Plasma protein binding and volume of distribution: determination, prediction and use in early drug discovery. In *Hit and Lead Profiling*, Faller B, Urban L Eds., WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Lombardo F, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, Harriman SP, Berellini G, Rajlic I, Obach RS (2013) Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 1: volume of distribution at steady state. *J Clin Pharmacol* **53:** 167-177.

Lombardo F, Jing Y (2016) In silico prediction of volume of distribution in humans. extensive data set and the exploration of linear and nonlinear methods coupled with molecular interaction fields descriptors. *J Chem Inf Model* **56:** 2042−2052.

Lombardo F, Berellini G, Obach RS, (2018) Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 1352 drug compounds. *Drug Metab Dispos* **47:**1466–1477.

Obach RS, Baxter JG, Liston TE, Silber BM, Jones BC, MacIntyre F, Rance DJ, Wastall P (1997) The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. *J Pharmacol Exp Ther* **283:**46 –58.

Obach RS, Lombardo F, Waters NJ (2008) Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 670 drug compounds. *Drug Metab Dispos* **36:**1385–1405.

Øie S, Tozer TN (1979) Effect of altered plasma protein binding on apparent volume of distribution. *J Pharm Sci* **68:**1203–1205.

Petersson C, Papasouliotis O, Lecomte M, Badolo L, Dolgos H (2019) Prediction of volume of distribution in humans: analysis of eight methods and their application in drug discovery. *Xenobiot* Published ahead of print. doi: 10.1080/00498254.2019.1625084.

Shimizu H, Yoshida K, Nakada T, Kojima K, Ogasawara A, Nakamaru Y, Yamazaki, H (2019) Prediction of human distribution volumes of compounds in various elimination phases using physiologically based pharmacokinetic modeling and experimental pharmacokinetics in animals. *Drug Metab Dispos* **47:**114–123.

Sui X, Sun J, Li H, Wang Y, Liu J, Liu X, Zhang W, Cheng L, He Z (2009) Prediction of volume of distribution values in human using immobilized artificial membrane partitioning coefficients, the fraction of compound ionized and plasma protein binding data. *Eur J Med Chem* **44:** 4455-4460.

Smith DA, Beaumont K, Maurer TS, Di L (2015) Volume of Distribution in Drug Design. *J Med Chem* **58:**5691–5698.

Ward KW, Smith BR (2004) A comprehensive quantitative and qualitative evaluation of extrapolation of intravenous pharmacokinetic parameters from rat, dog, and monkey to humans. II. Volume of distribution and mean residence time. *Drug Metab Dispos* **32:**612–619.

Waters NJ, Jones R, Williams G, Sohal B (2008) Validation of a rapid equilibrium dialysis approach for the measurement of plasma protein binding. *J Pharm Sci* **97:** 4586-4595.

Waters NJ, Lombardo F. (2010) Use of the Øie-Tozer model in understanding mechanisms and determinants of drug distribution. *Drug Metab Dispos* **38:**1159– 1165.

Figures legends.

Figure 1. Observed vs. predicted VD_{ss} values for the 191 compounds from Model 1. Dots are shaded according to their predicted f_{ut} values. The dashed, dotted and solid lines represent the line of unity, the 2-fold and 3-fold intervals, respectively.

Figure 2. Observed vs. predicted VD_{ss} values for the 191 compounds from Model 1. Dots are colored according to their charge class. The dashed, dotted and solid lines represent the line of unity, the 2-fold and 3-fold intervals, respectively. The vertical red line is set at 0.7 L/kg or total body water, while the blue vertical line is set at 2.8 L/kg.

Figure 3. Plot of the 191 compounds in the training set of Model 1, colored by fold-error and plotted in 3 dimensions using experimental $logD_{7.4}$ and f_{up} value as well as fold-error. Green dots represent compounds < 2-fold error, while orange and red dots represent compounds within 2- to 3-fold error and compounds > 3-fold-error, respectively. An increase in the number of red and orange compounds is noticeable at high $logD_{7.4}$and low f_{up}.

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 1. Coefficients and statistical parameters of multiple regression model based on 191 compounds using experimentally determined values. (Model 1, compound with f_{ut} < 0 were excluded)

Parameter	Value	Std. dev.	t-value	P > |t|
$logD^{7.4}$	-0.249	0.019	-12.975	< 0.0001
$f_{i(7.4)}$ (cationic)	-0.999	0.075	-13.342	< 0.0001
log f_{up}	0.735	0.053	13.841	< 0.0001
Intercept	0.070	0.072	0.970	0.333

TABLE 2. Coefficients and statistical parameters of multiple regression model based on 178 compounds using experimentally determined values. (Model 2, 13 compounds with $f_{ut} < 0$ and $f_{ut} > 1$ were excluded)

Parameter	Value	Std. dev.	t-value	P > |t|
$logD_{7.4}$	-0.231	0.018	-12.992	< 0.0001
$f_{i(7.4)}$ (cationic)	-0.929	0.070	-13.337	< 0.0001
log f_{up}	0.721	0.049	14.791	< 0.0001
Intercept	-0.040	0.070	-0.571	0.569

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 3. Statistics for the prediction of VD_{ss} and f_{ut} using Model 1, 2 and 3 (training set).

	Model 1 (N=191)		Model 2 (N=178)		Model 3 (N=176)	
Parameter	f_{ut}	VD_{ss}	f_{ut}	VD_{ss}	f_{ut}	VD_{ss}
GMFE	2.15	1.73	2.02	1.73	2.1	1.73
% within 2-fold (VD_{ss})	68		69		68	
% within 3-fold (VD_{ss})	87		86		88	
BIAS (obs-pred VD_{ss})	1.27		1.38		1.16	
Largest fold-error (VD_{ss}) (compound)	9.6 (colchicine)		10.1 (tebufelone)		9.2 (colchicine)	

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 4. Statistics for the prediction of VD_{ss} and f_{ut} using Model 1, 2 and 3 for the excluded sets.

	Model 1 (N=8, $f_{ut} < 0$)		Model 2 (N=13, $f_{ut} > 1$)		Model 3 (N=15, $f_{up} < 0.01$)	
Parameter	f_{ut}	VD_{ss}	f_{ut}	VD_{ss}	f_{ut}	VD_{ss}
GMFE	-	3.04	5.6	1.95	3.3	2.20
% within 2-fold (VD_{ss})	12		46		73	
% within 3-fold (VD_{ss})	62		85		73	
BIAS (obs-pred VD_{ss})	-0.18		-0.23		4.97	
Largest fold-error (VD_{ss}) (compound)	6.40 (glyburide)		3.64 (cephradine)		12.05 (tebufelone)	

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 5. Predictive performance of several classes of analogues using models built with the exclusion of each class of compounds (total N=191) and experimental descriptors.

Class	N	GMFE	% < 2-fold	% < 3-fold	Largest FE	Largest FE Compound
steroids	10	1.52	80	90	5.2	Ethinylestradiol
adrenergic	16	1.45	81	100	2.7	Nebivolol
NSAIDs	18	1.88	56	83	4.7	Suprofen
tri- and tetracyclic antidepressants	10	2.52	40	70	8.4	Maprotiline
benzodiazepines	12	1.61	75	92	4.1	Chlordiazepoxide
β-lactams	13	1.70	62	92	3.4	Cephradine
fluoroquinolones	11	1.42	82	91	3.3	Trovafloxacin
OVERALL GMFE	**90**	**1.69**	**68**	**89**	**8.4**	**Maprotiline**

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 6. Coefficients and statistical parameters of multiple regression model based on 191 compounds using computed $logD_{7.4}$ and $f_{i\,(7.4)}$. (Model 1c, compound with $f_{ut} < 0$ were excluded)

Parameter	Value	Std. dev.	t-value	P > |t|
$clogD_{7.4}$	-0.204	0.019	-10.756	< 0.0001
$cf_{i(7.4)}$ (cationic)	-0.924	0.084	-10.991	< 0.0001
log f_{up}	0.788	0.058	13.549	< 0.0001
Intercept	0.041	0.082	0.500	0.617

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 7. Predictive performance of several classes of analogues using models built with the exclusion of each class of compounds (total N=191) and computed $logD^{7.4}$ and $f_{i\ (7.4)}$ descriptors.

Class	N	GMFE	% < 2-fold	% < 3-fold	Largest FE	Largest FE Compound
steroids	10	1.70	80	90	5.3	Ethinylestradiol
adrenergic	16	1.50	81	94	3.1	Nebivolol
NSAIDs	18	2.07	56	78	6.1	Suprofen
tri- and tetracyclic antidepressants	10	2.73	40	70	11.0	Maprotiline
benzodiazepines	12	1.43	92	100	2.0	Delorazepam
β-lactams	13	2.22	46	62	5.6	Ampicillin
fluoroquinolones	11	1.38	91	100	2.1	Moxifloxacin
OVERALL GMFE	**90**	**1.81**	**69**	**84**	**11.0**	**Maprotiline**

DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Table 8. Performance of ionization classes based on recalculated VD_{ss} values (Model 1).

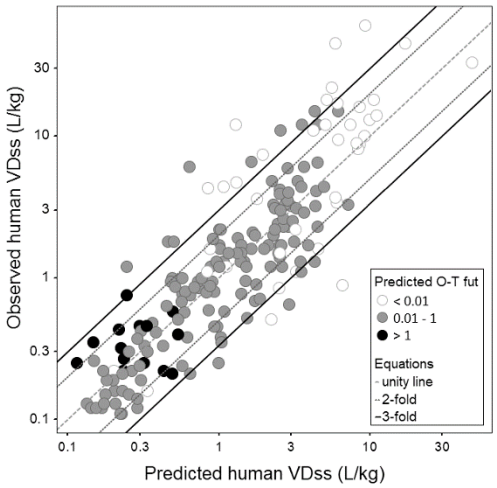
ionization class	N	GMFE	% < 2-fold	% < 3-fold	Largest FE	Largest FE Compound
neutral	53	1.88	62	79	9.6	Colchicine
cationic	70	1.82	66	84	7.6	Maprotiline
anionic	47	1.49	79	96	3.1	Cephradine
zwitterionic	21	1.65	67	95	3.3	Oxytetracycline

Table 9. Predictive performance of Model 1 across predicted VD_{ss} ranges.

Range of predicted VD_{ss} (L/kg)	N	GMFE	% < 2-fold	% < 3-fold	BIAS (obs-pred)	Largest FE	Largest FE Compound
<0.7	65	1.59	75	91	0.16	9.60	Colchicine
0.7 - 2.8	74	1.71	69	88	0.50	9.27	Tebufelone
>2.8	52	1.96	58	81	3.76	7.64	Maprotiline

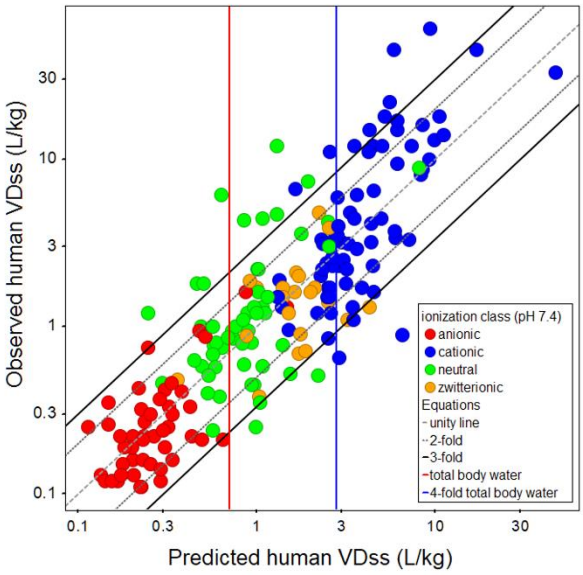
DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 1.



DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 2.



DMD Fast Forward. Published on October 2, 2019 as DOI: 10.1124/dmd.119.088914
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 3.