DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

TREND ANALYSIS OF A DATABASE OF INTRAVENOUS PHARMACOKINETIC PARAMETERS IN HUMANS FOR 1352 DRUG COMPOUNDS

Franco Lombardo, Giuliano Berellini, and R. Scott Obach

Drug Metabolism and Bioanalysis Group, Alkermes Inc., Waltham, Massachusetts (F.L.); Computational Chemistry Group, Biogen, Inc., Cambridge, Massachusetts (G.B.); and Pharmacokinetics Dynamics and Metabolism Department, Pfizer Global Research and Development, Groton Laboratories, Groton, Connecticut (R.S.O.)

Running Title: Trend analysis of human intravenous pharmacokinetic parameters.

Corresponding author: 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: 27

Number of tables: 1

Number of figures: 9

Number of references: 47

Abstract: 206 words

Introduction: 1055 words

Discussion: 1914 words

Non-standard abbreviations:
ADME = Absorption, Distribution, Metabolism, Excretion, ADMET = Absorption, Distribution, Metabolism, Excretion, Toxicity, CAS = Chemical Abstracts Service, MW = molecular weight, PSA = Polar Surface area, SMILES = simplified molecular-input line-entry system

Abstract

We report a trend analysis of human intravenous pharmacokinetic data on a data set of 1352 drugs. The aim in building this dataset and its detailed analysis was to provide, as in the previous case published in 2008, an extended, robust and accurate resource which could be applied by drug metabolism, clinical pharmacology and medicinal chemistry scientists to a variety of scaling approaches. All in vivo data were obtained or derived from original references, either through literature or regulatory agencies reports, exclusively from studies utilizing intravenous administration. Plasma protein binding data were collected from other available sources to supplement these pharmacokinetic data. These parameters were analyzed concurrently with a range of physicochemical properties and resultant trends and patterns within the data are presented. In addition, the date of first disclosure of each molecule was reported and the potential "temporal" impact on data trends was analyzed. The findings reported here are consistent with earlier described trends between pharmacokinetic behavior and physicochemical properties. Furthermore, the availability of a large data set of pharmacokinetic data in human will be important to further pursue analyses of physicochemical properties, trends and modelling efforts, and should propel our deeper understanding, especially in terms of clearance, of the ADME behaviour of drug compounds.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

The last twenty years or so have seen the flourishing of prediction approaches of human pharmacokinetics for new compounds and very recent work, in the form of perspective articles, on in silico ADME (Lombardo, 2017) and several modeling endeavors (Berellini, 2012; Gombar 2013; Lombardo, 2016), highlight that observation. Many reports, of which we cite a few examples, have focused on scaling techniques, i.e. approaches and techniques that utilize animal pharmacokinetic data (Ward 2004a, 2004b; Caldwell, 2004; Jolivette, 2005; Evans, 2006; Mahmood, 2006; Martinez, 2006; Tang, 2006; McGinnity, 2007; Fagerholm, 2007; Lombardo, 2013a and 2013b) as well as in vitro data (Obach, 1997; Lombardo, 2002; Nestorov, 2002; Lombardo, 2004; Riley, 2005; Grime, 2006). The growing availability and acceptance of computational chemistry methodologies has generated many examples of prediction of human pharmacokinetics and/or general ADMET properties (Veber, 2002; Cruciani, 2005; Gleeson, 2006; Gleeson, 2007; Ghafourian, 2006; Lombardo, 2006; Gunturi, 2007; Norinder, 2007; Berellini, 2009; Berellini, 2012; Lombardo, 2016, Lombardo, 2017). An essential pillar of these methods, beyond a careful choice of the nature and type of descriptors (e.g. 2D or 3D, fragmental or continuous) and the theory behind statistical approaches, is the availability of curated data and relatively large databases that have been carefully assembled. The difficulties in building a human pharmacokinetics database are many-fold. At the onset there is the difficulty of finding a vast array of intravenous data within any single company and each publically available study (each yielding a data point) is diverse and separate in time, as well as in the experimental

approach(es) taken by different investigators. The reader will easily recognize, among the variables, the number and types of study subjects (e.g. healthy vs. diseased, gender, age, etc), the routes of administration (all intravenous in the present case) and doses, sample collection times, methods of analysis of the samples, and the types of pharmacokinetic parameters reported. This database spans many decades and the improvement in technology and sensitivity of analytical methods, as well as our understanding of pharmacokinetics, clinical trials design and other trends in related sciences, all have influenced and shaped the data available. We discuss the data assembly process and the pharmacokinetic parameters we have evaluated in their collection, in detail, in the Materials and Methods section.

We have defined above the availability of quality data as a "pillar" of modeling work. A key aspect in the development of models for the prediction of human clearance (CL), volume of distribution (VD), and absolute oral bioavailability, is that data used are obtained from studies based on intravenously administered doses. We also note here that definitions vary and the methods of calculation adopted need be consistent. We offer, as an example, the "variable" nature of volume of distribution, which may be taken as steady-state VD (indicated as V_{ss} or VD_{ss}), central VD (generally indicated as V_c or VD_c) or terminal phase VD (generally indicated as VD_β or VD_z, with or without the letter D). This is another aspect we have concentrated our attention toward, as in the previous report, and we discuss it in the Materials and Methods. As an example, that the extensive set of human pharmacokinetic data reported in Appendix II of "The Pharmacological Basis of Therapeutics", (Goodman and Gilman's, 13th Ed., 2017) is frequently cited as a source of data for model construction. This data set, highly curated by many successive

groups of very experienced scientists in the field, was not intended for the development of structure-pharmacokinetic relationships. It aimed, instead, at offering a guide for dosing regimens through understanding of pharmacokinetics basis to health care professionals and medical students. As such, and in many cases, the data have been reported using oral administration with volume parameters that include terminal phase data and not only VD_{ss}. The performance of models may likely be confounded if this set is utilized.

The objectives of this study were 3-fold and we sought: 1) to greatly expand on the already existing large publicly available database we reported earlier (Obach, 2008) of carefully scrutinized human pharmacokinetic parameters, which would be easy to consult and reference, and which could be used by scientists at early stages of drug research for the construction of predictive pharmacokinetics models, 2) to gain some insight into the relationships between chemical properties, as derived from structural attributes (i.e. computed descriptors), and human pharmacokinetic parameters and 3) to try to infer, with the addition of "time data" represented by the year of first disclosure, whether "temporal trends" and/or changes could be discerned across the many decades of drug and pharmacokinetic research. This is similar to the development of several notable guidelines, such as the rule of 5 (Lipinski, 1997) or the rotatable bond limits described by Veber (Veber, 2002) stemming from similar observations, although the aim of both reports was to derive guidelines toward improved oral absorption.

In pursuit of the first objective, we have carefully examined a vast set of scientific literature, to very recent publications, for the human pharmacokinetic parameters CL_p, (clearance in plasma) $t_{1/2}$, VD_{ss}, and MRT measured after intravenous administration. We

were able to expand the data set by successfully obtaining human intravenous pharmacokinetic data as to double the original work. Plasma protein binding data are included for the majority of these compounds, albeit not for all of them, despite very extensive data searches through multiple sources. This database, now comprised of 1352 compounds, is provided as a data table and as an Excel spreadsheet with CAS number, SMILES, original reference and all the computed parameters utilized in this work, as supplemental information accessible to all readers. There are no biologics, large proteins or monoclonal antibodies in the present or previous data set, but several compounds, represented largely by therapeutic peptides, have a MW in excess of 600 Da and up to 7150 Da (mipomersen). The database should offer a solid starting point to scientists aiming at developing correlative analysis and/or computational models toward the prediction of human pharmacokinetics of new compounds. Furthermore, using this database, we have examined the relationships between human intravenous pharmacokinetic parameters ($VD_{ss,u}$, VD_{ss}, $CL_{p,u}$, CL_p, MRT, f_u and $t_{1/2}$) and various computed fundamental physico-chemical parameters (e.g. logD, charge, polar surface area, etc.) and offered initial and general consideration on its span and median values. Here f_u denotes the fraction unbound in plasma and the subscript "u" has the usual meaning of unbound.

Materials and Methods

Data mining. There are, as readers will recognize, several ways to report data which often are without specification (as mentioned above for volume of distribution) of the modality of calculation. This leads to the use of various units and symbols, whether a compartmental analysis is adopted vs. a non-compartmental approach. Therefore, the description and details concerning methods, reported by authors can vary considerably. The examination of each individual report and the extraction of fundamental pharmacokinetic parameters required careful scrutiny. The result of these efforts, however, has been the extension of the original dataset to more than a doubled number of compounds from our previous publication (Obach, 2008). The present work was conducted using the same criteria adopted for the previous paper which may be consulted for a detailed account of calculation methods and equations used. In addition, we have searched SciFinder® for the date of the first published report for each molecule. The dates entered may or may not reflect the actual discovery, for example because disclosure in a patent may have followed a few years after synthesis or isolation, nor they might always be expected to represent the introduction into therapy, which may have followed at a significantly later point. However, using a general scheme of "binning" by decade, we sought potentially useful and informative temporal trends. As in the previous case, we did not include any data from oral, intramuscular, or any other dosing route wherein the total dose may not have entered the systemic circulation, since the calculation of pharmacokinetic parameters CL and VD require the dose available to the central (plasma) compartment. Intravenous data were either from rapid bolus injection or infusions. In addition, we tried to carefully identify compounds, or dose ranges, where non-linearity

was reported, and utilized the lowest dose or lowest range where linearity was observed and good analytical data was reported. We did not check upon nor did we consider the relatively rare event of reversible metabolism (e.g. N-oxide formation and reduction) having a significant impact on the overall values for any of the compounds, and used the data as the average values reported by the authors.

Careful evaluation of VD_{ss} calculations reported, this being a volume term which is more generally related to an overall distributional behavior, was performed along the same lines described in the equations detailed in the previous work, and we refer the reader to that work. As an additional source, prescribing information and/or biopharmaceutics reviews available on-line through the FDA website, in which the data listed have been reviewed and approved by the U.S. Food and Drug Administration, were sometimes the only source of VD_{ss} data. As in the previous work, a 70 kg average weight was assumed if not reported or a mid-point of the range given was taken. Similarly, when data were reported using BSA we utilized in all case a 70 kg value against a body surface area of 1.73 m^2. We also extensively searched for protein binding values in human plasma (or serum). The notes and comments in the full set of data, available as supplemental information, offer more specific indication on particular compounds where, for example, digitization of concentration vs. time plots was performed to calculate otherwise unavailable PK parameters.

Some of the physicochemical parameters calculated for all compounds and described in the Results section were calculated via the MoKa software (v.2.6.6; Molecular Discovery, Hertfordshire, UK). They are logP and $logD_{7.4}$ as well as the pK_a values to determine the ionization state, the latter assigned on the basis of the most abundant species at pH 7.4. In a few cases (15 out of 1352 compounds) some compounds

did not yield these physicochemical descriptors we sought to calculate either because of relatively large molecules (e.g. mipomersen, MW 7150 Da) or because of the presence of metal ions (e.g. Gd or Pt) in the compounds. These compounds are included in the data table of PK parameters and general characteristics but omitted from descriptors statistics considerations. The number of rotatable bonds (NRB), the number of hydrogen bond donor and acceptor atoms (HBD and HBA) and N and O atoms polar surface area (in $Å^2$), were calculated via Vortex (v2017.08.69034.51-s).

RESULTS

Characteristics of the Pharmacokinetic Values. As described in the Methods section, we undertook extensive mining of the scientific literature. The reanalysis of concentration vs. time data was performed in some cases, allowing the collection of human intravenous pharmacokinetic parameters for a total of 1352 compounds. In some cases we could not find a complete set of pharmacokinetic parameters but the compound was included as there was either clearance or VD_{ss} and we deemed the information useful. In addition, for for many of the compounds (920/1352), we were able to find plasma protein binding data. These parameters are reported as a table in the supporting information and as a spreadsheet containing all the values along with literature references, comments and notes, and both files are included as attachments in the Data Supplement. The data span over considerable ranges (Figure 1). Secretin (a peptide) had the lowest VD_{ss} value at 0.03 L/kg, while hydroxychloroquine had the highest value at 700 L/kg. As previously observed the vast majority (1171/1315 data points or 89%) were comprised within 0.1 and 10 L/kg, as shown in Table 1. The mean and median values were 3.8 and 0.9 L/kg, across all data, or essentially identical to the values observed in the previous work (4.2 and 0.96 L/kg, respectively). Forty-three percent of the compounds (561/1315) have VD_{ss} values at or below 0.7 L/kg (Table 1) generally considered, to be the value for total body water, and the percentage is essentially identical to the value reported for the set of 670 compounds (41%). Lastly, eight percent of compounds (108/1315) had VD_{ss} values of or greater than 10 L/kg, an indication of extensive partitioning into tissues, a percentage identical to the previously reported value. Thus the general statistics and

characteristics of the VD_{ss} values are the same as in the previous report from 10 years ago.

Plasma clearance values (Figure 1B) ranged from 7-hydroxystaurosporine, in this set, represents the compound with the lowest clearance at 0.004 mL/min/kg while, on the other end of the range, the highest value belongs to artesunate with a value of 1070 mL/min/kg, and these limits are held by the same two compounds as in the previous work, with mean and median values of 12.2 and 4.5 mL/min/kg. Both values did increase slightly from values of 10 and 4 mL/min/kg, in the previous work. A very similar proportion to what was previously observed, of sixty-eight percent of the compounds resided in a range between 1 and 15 mL/min/kg (919/1350) as shown in Table 1, with sixteen percent (215/1350) possessing clearance values below 1 mL/min/kg (very low clearance), which is identical to the value of 16% reported for the previous set. In this set 135 of the compounds had CL values greater than liver blood flow (Table 1) taken as 21 mL/min/kg, as opposed to 56 compounds in the previous set which is slightly more than double the number of the compounds. However, considering the doubling of the data set, this cannot be construed as an indication of a trend toward a higher "acceptance" and progression to clinical studies of high clearance compounds. Regardless of the statistical implications, these values are suggestive of the possibility of blood-to-plasma ratios greater than unity or of extrahepatic clearance mechanisms. Several compounds possessing high CL values can be classified among anesthetics, systemic and short-acting such as propofol, local such as articaine or prilocaine, among pain medications such as remifentanil, butorphanol, dezocine or pentazocine, or among cytotoxic cancer chemotherapeutics, such as amifostine, laromustine or carmustine, which are drug classes that are frequently administered via the intravenous route to optimize therapy or use. The

latter few compounds exert their action by acute cytotoxicity, which would be dangerous if prolonged and thus the short half-life is beneficial toward controlling side effects. Prodrugs, on the other hand, are by definition compounds in which a pharmacologically active metabolite is being formed from the parent and thus are expected, and desired to have, high CL (e.g. dolasetron, esmolol, etc) to yield, as rapidly and completely as possible, the active moiety.

Terminal phase half-life values ranged from 1.2 min (perflutren) to 56 days (almitrin) as shown in Figure 1C, with sixty-three percent (843/1335) residing between 1 and 12 h (Table 1). The bisphosphonates may be prone to underestimated half-lives since they sequester into bone and, by doing so, would not be detectable in blood or plasma. The average $t_{1/2}$ was 17.1 h and the median value was 4.5 h. Half-life values, being a derived parameter, do not lend themselves to a direct correlation with physiological properties represented by volume of total body water or hepatic or renal blood flow. However they can, if there is no considerable difference between pharmacodynamics and pharmacokinetics, be classified into ranges largely based on dosing frequency values. In pharmaceutical research, scientists very often seek therapeutics which are amenable to q.d. dosing, being considered convenient from a patient compliance standpoint. Looking at $t_{1/2}$ (or MRT) and using the above requirement approximately about three-fourths of the compounds in the dataset (1015/1335) have $t_{1/2}$ values below 12 h (Table 1). Therefore they would likely require a more frequent dosing regimen. Plotting MRT vs. clearance (plot not shown), we estimated that a value of < 1 mL/min/kg for clearance would be needed, to achieve at least a MRT value of 4 hours for acidic compounds, because acidic compounds (with significant exceptions) yield a generally lower VD_{ss}.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

The range and distribution of plasma protein binding values is in Figure 1D and Table 1. The values ranged between no binding (several compounds) and 0.0002 fraction unbound (amiodarone), with mean and median free fractions of 0.35 and 0.2. Only sixty percent (560/900) of the compounds in the set have f_u values greater than 0.1, and about twelve percent, or one in eight compounds (111/920), could be considered highly bound ($f_u < 0.01$; Table 1).

Characteristics of the Computed Physicochemical Values. The 1352 compounds in this dataset span a wide range of fundamental computed physicochemical characteristics (Figure 2). The typical drug-like space for molecular weight (200 to 600 Da) is illustrated in Figure 2A and it is covered by seventy-eight percent of the compounds (1060/1352), with median value of 371 Da and a range from 42 (cyanamide) to 7150 Da (mipomersen). Furthermore, the median and mean number of rotatable bonds (NRB), was calculated as 8 and 5, and the median and mean value of calculated PSA were 89 and 130 $Å^2$ (PSA counting nitrogen and oxygen atoms only), and they are below the upper limit reported (10 for NRB and 140 $Å^2$ for PSA) by Veber (Veber, 2002), for compounds with good probability of being orally bioavailable in rat. The binning is illustrated by Figure 2B and Figure 2C for NRB and PSA, respectively.

There were, among the compounds for which the pKa could be calculated (some large peptides and ion-containing compounds were not amenable to computation) 313 anionic, 472 cationic, 457 neutral and 97 zwitterionic compounds (Figure 2D). These were categorized by calculating the most abundant species (anionic for acids, cationic for bases, both for zwitterions) at pH 7.4, using the calculated pK_a values from MoKa.

Thus zwitterionic compounds represent a percentage of compounds well below 10% of the entire dataset. The same was true for the previous data set (Obach, 2008) but,

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

while the basic and acidic compounds essentially doubled (from 267 and 159, respectively) and neutral compounds almost tripled (from 173), zwitterions increased only by 50%.

Lipophilicity, expressed as clogP and $clogD_{7.4}$, yielded 2 and 0.7 as median values while the corresponding mean values were 1.6 for ClogP and 0.1 for $clogD_{7.4}$. Figures 2E and 2F show the bins and number of compounds for clogP and $clogD_{7.4}$. The median and mean values for numbers of hydrogen bond acceptors were 6 and 9, while the corresponding values for hydrogen bond donors were 2 and 3 to 4, and the binning is shown in Figure 2G and 2H. The median and mean clogP values, and similarly the number of hydrogen bond acceptors and donors, although coming from compounds dosed intravenously in all cases, were well below the limits set (5, 10 and 5 respectively) by the well-known Lipinski "rule of 5" (Lipinski, 1997). It could be argued that the "rule of 5"sets a threshold for potential issues with oral absorption, while these compounds were all dosed intravenously. However, it is true that most of them were chosen for development as oral drugs, for which intravenous data were also generated.

Trends in the Dataset: VD_{ss} vs. Physicochemical Properties. The dataset was mined for discernible trends between the computed physicochemical properties and VD_{ss} values in human. No single physicochemical descriptor (property) yielded a relationship that could be, on its own, predictive of VD_{ss}. However, as observed in the past, "composite" contributions to VD_{ss} by various physiochemical properties could be observed through trends in the data. Overlap and scatter in the data, however, made observation of trends difficult when values were partitioned by means, (as exemplified in Figure 3 A-F), but when using median values they could be observed. VD_{ss} values yielded observable trends with $clogD_{7.4}$, PSA, and number of H-bond acceptors and donors (Figures 3B-C,

3E-F), as well as charge type (Figure 3A), which have varying degrees of relatedness to each other. High PSA, high numbers of H-bond acceptors/donors, and low lipophilicity offer recognizable trends: median VD_{ss} values trend higher for low PSA, higher for low numbers of H-bond acceptors/donors, and high with lipophilicity (respective panels in Figure 3). Acidic compounds (anions) yield generally lower median VD_{ss} values than zwitterionic and neutral compounds, which are, in turn, lower than basic compounds (Figure 3A). However, there is a very large overlap and the trend can only be taken as fairly broad, as it is possible to encounter many acidic compounds that show a fairly sizable VD_{ss} value, often well above total body water, taken as reference at a value of approximately 0.7 L/kg. A median value of around 0.2 L/kg in VD_{ss} for the acids was observed, but an upward trend was discernible, for free VD_{ss} with increasing $logD_{7.4}$ for bases, neutrals and zwitterions (see Figure 6B).

Trends in the Dataset: CL vs Physicochemical Properties. Discernible trends, as in the case of VD_{ss}, could be readily observed between clearance and some of the physicochemical properties. However the quantitative prediction of clearance on the basis of any single property was not supported by our observations, since we could not find a strong enough relationship allowing such prediction. Decreases in median CL were observed with increases in PSA or with an increase in hydrogen bond acceptors and donors (Figure 4C and 4E-F). Only a weak trend could be observed between median CL and lipophilicity. Free clearance showed generally a more discernible trend toward higher values for basic and neutral compounds values than for acids or zwitterions (Figure 6D, lower right panel) although data overlap to a great extent.

Trends in the Dataset: Protein Binding vs Physicochemical Properties. Two relationships, one between protein binding and lipophilicity and one between binding and

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

charge class, could be discerned (Figure 5A-B). Lipophilicity shows an increasing (direct) trend across all charge types. Another observation, as it may be expected, is a discernible lower median f_u value for anionic (acidic) compounds vs. basic and neutral compounds, and the much higher median f_u value for zwitterionic compounds. This may be due to their amphiprotic nature resulting in a possible lower affinity for albumin or α1-acid glycoprotein, as well as other plasma proteins. However, as shown by Figure 5A, there is a great deal of overlap among all classes, due to a wide distribution and, contrary to the accepted perception of a significantly lower f_u for acidic compounds, this cannot be generalized.

Trends in the Dataset: free VD_{ss} and CL vs. $logD_{7.4}$. We examined VD_{ss} data after correcting for free fraction for all compounds with available f_u, (i.e. free VD_{ss}= VD_{ss}/f_u), and this transformation greatly expands the range of values, together with offering a perhaps better indication of the extent of tissue binding (Figure 6A and 6B).

Total and free VD_{ss} were plotted vs. $logD_{7.4}$ ($logD_{7.4}$ range -6 to 6). There is no apparent trend using total values (Figure 6A, red dots) for acidic compounds, while an upward trend is recognizable for neutral and basic compounds. Once the values are corrected via f_u, a clearer trend is apparent (Figure 6B) which reveals a shift toward higher values of free VD_{ss} with $logD_{7.4}$ and more apparent for acidic compounds. The data is therefore suggestive of plasma protein binding having the tendency to dominate the distribution behavior of negatively charged compounds.

CL values were similarly examined before and after correction for f_u (i.e. free CL = CL/f_u) as shown Figure 6C-D. As in the case of VD_{ss} correction, the data sets are smaller reflecting the lower number of compounds for which we found f_u values. At first glance the effect seems to parallel the trend observed for VD_{ss} but the profile seems

flatter for clearance using total values, than in the case of total VD_{ss}. However, the potential involvement of uptake and efflux transporters notwithstanding (Waters, 2010), it should be borne in mind that VD_{ss} is generally, dominated by physicochemical properties, largely fraction ionized and lipophilicity, even after removal of the protein binding aspect (Lombardo, 2002; Lombardo, 2004). As opposed to VD, CL is largely dependent on affinities and intrinsic activities for specific enzymes and transporters. Affinities and intrinsic activities will be somewhat dependent on basic physicochemical properties, but also on the interactions of specific substituents and fragments with macromolecules involved in drug metabolism and disposition.

Trends in the Dataset: Time-dependent Variations. A plot of year of first appearance, binned according to the year and colored by VD_{ss} (Figure 7, left panel), CL (Figure 7, right panel), fraction unbound in plasma (Figure 8, left panel) and $clogD_{7.4}$ (Figure 8, right panel) does not seem to indicate any particular trend for the first two properties (using the same number of compounds with f_u data available) while some increase toward more lipophilic compound and lower f_u (as they are generally inversely correlated) could be discerned, as shown in Figure 8. In particular, the bin representing compounds $clogD_{7.4}$ between -1 and 0 (Figure 8, right panel) was reduced to a very low 4% in the period 2000 to present, possibly influenced by the significantly fewer number of total compounds we could retrieve from the literature.

A plot of year of first appearance in the literature, (Figure 9) binned as approximately two decades per section, and with compounds reported prior to 1960 as one bin, shows a significant trend with compounds having a MW above the median value of the present set (371 Da) increasing steadily. The rate is 20% above the previous twenty years starting from 1960-1980 and reaching 80% in the period from 2000 to

present. This finding was also coupled with a significant reduction in compounds reported, which showed a decrease by a 3-fold margin from the two decades spanning 1980-2000.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

DISCUSSION

In this work, we sought to expand the human pharmacokinetic dataset that we originally described in 2008 (Obach, 2008). The number of compounds in the current set is approximately doubled, and thus merited a renewed evaluation of the overall trends and relationships between the pharmacokinetic parameters and fundamental physicochemical properties. The set of 1352 compounds encompasses, as before, a wide variety of drugs in a broad range of therapeutic areas and, consequently, a wide variety of structural characteristic, pharmacokinetic values and physicochemical descriptors. The data were carefully curated (as described in detail in the Methods section) and strictly from intravenous administration. Thus, these data should not only be of use for the present analysis, but they are available (in the Supplementary Information) for others to utilize to develop other relationships and models. Overall, the trends in the pharmacokinetic parameters between the original dataset and the present doubled dataset were the same; ranges, means, and medians were largely unchanged. In the present work we also examined the impact of the year of first disclosure, as reported in SciFinder® although those data do not necessarily reflect the year of discovery (likely earlier) or the introduction of the compound into therapy, which may have happened at a significant later point. It is possible that, by binning the time ranges (see Results) by two decades (or overall before 1960), we may have attenuated such differences and grouped compounds in reasonably comparable "periods". We will discuss time-related findings below.

Among the pharmacokinetic parameters collected, volume of distribution at steady state (VD_{ss}) is the one that has the most marked relationships with physicochemical properties. VD is largely a function of differential partitioning between plasma and other tissues, which in turn is a function of non-specific binding to tissue

components and plasma proteins, such as albumin. Such non-specific interactions are largely dependent on the physicochemical characteristics of the drug. Charge state has an influence with cationic compounds showing generally greater VD_{ss} values, but there is considerable overlap which shows that other factors such as lipophilicity also have an influence (Figure 3). The relationship to lipophilicity becomes stronger when VD_{ss} is corrected for plasma free fraction (Figure 6). VD_{ss} shows the same proportion of compounds (43%) with value < 0.7 L/kg, taken as total body water, while only a small proportion is confined to blood volume (taken as 0.1 L/kg). Therefore, a large proportion does exceed the total body water value, but this is only an indication of the resulting "distributional average" of the compound in the body and it does not inform about the presence of the compound in a particular organ or at an intended target.

In Figure 6 we show the trend and correlation between computed $logD_{7.4}$ with either VD_{ss} or clearance both total (panels A-B) and free (panels C-D) and we note that, while there is a seemingly clearer trend in the case of the free value of both parameters, the one for VD_{ss} starts being more noticeable even when examined via total VD_{ss} while the corresponding plot for total clearance does not allow discernment of much of a trend. In fact, Figure 3B (median and distribution of total VD_{ss}) does show an increment of the median value for total VD_{ss} vs. computed $logD_{7.4}$ while the corresponding plot (Figure 4B) for total clearance does not. If we compare more broadly Figure 3 (B-F panels) for VD_{ss} and the corresponding Figure 4 (B-F panels) for clearance, we note that not much of a correlation is discernible between (total) clearance and the computed descriptors shown in the case of clearance. At the same time (Figure 3) VD_{ss} shows trends between median values and computed descriptors which are opposite, as expected, in the case of $logD_{7.4}$ vs. polar surface area as well as HBA and HBD. In the case of the correlation of free

parameters with computed $logD_{7.4}$ the VD_{ss} plot (Figure 6B) seems to yield a higher positive slope than does the plot for free clearance vs. computed $logD_{7.4}$ (Figure 6D).

Unlike VD, clearance is driven by interactions between drugs and the drug metabolizing enzymes and/or drug transporters involved in their clearance, as well as plasma protein binding. The interaction of individual drugs with enzymes and transporters is more a function of specific ligand-protein interactions as opposed to non-specific interactions, thus relationships of between gross physicochemical properties and clearance should not be as apparent as they are for VD (see Figure 4). Compounds with low free fractions could also have lower CL, so after correction of CL to free CL, a slight relationship between free CL and lipophilicity can be observed (Figure 6). But overall, relationships between free CL and physicochemical parameters are not nearly as discernible as they are for free VD. In addition, investigators have reported computational models where they utilized continuous physicochemical descriptors for models of volume of distribution (Berellini et al., 2009; Lombardo et al., 2016; Gleeson, 2006) while they found it necessary to use structural descriptors (i.e. fragments) to improve the predictive power of clearance models even though the prediction of the general clearance mechanism (metabolic vs. renal vs. biliary) did not require the latter descriptors for a good performance (Berellini et al., 2012; Lombardo et al., 2014). Therefore, lipophilicity alone does not describe the clearance behavior of drug compounds although it is certainly an important component for the elimination of xenobiotics via more polar and water-soluble compounds.

In general, the behavior of metabolic enzymes (and transporters) can be quite complex, and many examples can be found in the literature where identical lipophilicity yields a very different clearance outcome. Smith (Smith, 1997) pointed out that clearance

differences are related to the propensity toward N-demethylation in a small series of benzodiazepines rather than bulk lipophilicity. Similarly, Stepan described the discovery of a sizable series of γ-secretase inhibitors where, in many cases, substitution, regioisomerism and stereochemistry were responsible for large variations in scaled in vitro hepatic clearance while changes in experimental ElogD were barely discernible or not at all measurable (Stepan, 2011). The outcome, due to the complexity, redundancy and promiscuity of metabolic enzymes and transporters, layered upon selectivity and safety considerations, is clearly very complex and multidimensional. Along the same lines are the comments of other researchers (Broccatelli, 2018) stating that "t1/2 optimization via lipophilicity reduction without addressing a metabolic soft-spot is unlikely to work".

In addition to the pharmacokinetic parameters, we also collected plasma protein binding data for as many of the compounds in the dataset as available. This was done primarily in order to be able to correct total VD and CL values to free values for comparison to physicochemical properties. However, with these data, we also could compare free fraction values to physicochemical properties. Plasma binding is mostly driven by albumin (which is at ~ 600 μM or 42 g/L, Davies and Morris 1993) and α1-acid glycoprotein (at ~ 43 μM or 1.8 g/L, Davies and Morris 1993), with the former mostly associated with binding anionic drugs and the latter, primarily but not exclusively, associated with binding cationic drugs (Ghuman 2005, Meijer and Van der Sluijs 1987, Israili and Dayton 2001). Protein binding correlated with lipophilicity (Figure 5). We have also pointed out, in the Results section, that the generally accepted notion of a much higher binding for acidic compounds is not strongly supported by "clustering" for anionic, cationic and neutral molecules, although the anionic compounds, acidic in

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

nature, do seem to show a lower overall median f_u than basic (cationic) compounds. There is a great deal of overlap and many "exceptions" exist to the perceived much greater binding of anionic compounds to plasma proteins. Zwitterionic compounds do show a discernibly different median, likely due to their ability to interact with a broader set of proteins but, once more, with a great deal of overlap and not much clustering. Lipophilicity (Figure 5B) and number of HBD (Figure 5F) do show a negative (increasing $logD_{7.4}$ decreases f_u) and a positive (increasing number of HBD increases f_u) trend, respectively, although very few compounds are present in the 7-10 bin for HBD and a much lower median is observed for the uppermost bin, probably due to the presence of larger molecules and opposing factors. The flexibility of a molecule, expressed as the number of rotatable bonds (Figure 5D) does seem to show a negative correlation, perhaps only for molecules within the lower 3 bins, and flexibility may play a detrimental role toward free fraction, similar to the effect on absorption. However, this effect, if real, manifests itself well before the classical threshold outlined by Veber (Veber 2002) of 10 rotatable bonds. We also note that the original observation was reported with the aim of exploring the effect of flexibility on absorption, and the dataset was based on permeability across artificial membranes, so the significance in this context is not clear.

Finally, we also gathered data to indicate whether human pharmacokinetic, and even physicochemical properties, have been changing over time. We used the date of first disclosure of a compound which is not an entirely accurate description of when a compound was first synthesized or discovered, but it offered the best surrogate for the analysis. What is interesting is that while lipophilicity of the drugs in our dataset increases in more recent years (Figure 8, right panel), values of CL and VD have generally remained constant (Figure 7) albeit VD values have increased a bit. Another

observation we wish to highlight are the findings illustrated by Figure 9, that is, a significant trend toward higher molecular weight in more recent times. A possible explanation for this behavior may be represented by the exploration of different and more complex drug space such as, for example, protein-protein interaction (PPI) and, therefore, the pursuit of larger molecules needed to disrupt shallower PPI clefts. This may include exploration of peptide drugs with perhaps a significant deviation from the earlier oral drug paradigm, in order to achieve modulation of otherwise inaccessible therapeutic targets. Another possibility is represented by the expansion of techniques and trends in combinatorial chemistry which has influenced upward the molecular weight of compound libraries across industry perhaps most notably in the 1980s and 1990s. This may have manifested a bit later with larger compounds entering clinical trials. Increasing MW and lipohilicity in new drugs may be due to the desire to impart increased potency and/or greater target selectivity. Such desired properties can require the generation of larger, more lipophilic drugs. As mentioned above, increased lipophilicity can yield increased plasma protein binding, but also increased tissue binding and increased metabolic intrinsic clearance. These properties may all 'cancel' each other out and thereby CL and VD may not change, as shown in Figure 7.

In conclusion, we have summarized a human pharmacokinetic dataset that is the largest of its kind to our knowledge. We have exhaustively searched the scientific literature and other sources for bona fide human intravenous pharmacokinetic studies and scrutinized the methods and data presented. These data have proven valuable in examining relationships between fundamental human pharmacokinetic parameters VD_{ss} and CL with various basic physicochemical properties. The dataset in this report approximately doubles our previous report from a decade ago (Obach, 2008), yet the

relationships between human pharmacokinetic parameters and physicochemical properties have remained largely unchanged. These data, available in the Supplemental Information available on-line, can be utilized and mined by others interested in deriving relationships between structure and human pharmacokinetics. Our own efforts are ongoing to establish whether trends relating particular structural entities and human pharmacokinetic parameters can be determined.

Authorship contributions

Participated in research design: Berellini, Lombardo, Obach;

Performed data acquisition and search: Berellini, Lombardo, Obach;

Performed data analysis: Berellini, Lombardo, Obach;

Wrote or contributed to the writing of the manuscript: Berellini, Lombardo, Obach.

REFERENCES

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

Berellini G, Waters NJ, Lombardo F (2012) In silico prediction of total Human plasma clearance. *J Chem Inf Model* **52:** 2069–2078.

Broccatelli F, Aliagas I, Zheng H (2018) Why decreasing lipophilicity alone is often not a reliable strategy for extending iv half-life. *Med Chem Lett* ***9:*** 522-527.

Caldwell GW, Masucci JA, Yan Z, and Hageman W (2004) Allometric scaling of pharmacokinetic parameters in drug discovery: can human CL, Vss and t1/2 be predicted from in-vivo rat data? *Eur J Drug Metab Pharmacokin* **29:**133–143.

Cruciani G, Carosati E, De Boeck B, Ethirajulu K, Mackie C, Howe T, and Vianello R (2005) MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J Med Chem* **48:**6970 – 6979.

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

Davies B, and Morris T (1993) Physiological parameters in laboratory animals and humans. *Pharm Res* **10:** 1093-1095.

Evans CA, Jolivette LJ, Nagilla R, and Ward KW (2006) Extrapolation of preclinical pharma- cokinetics and molecular feature analysis of "discovery-like" molecules to predict human pharmacokinetics. *Drug Metab Dispos* **34:**1255–1265.

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

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

Ghuman J, Zunszain PA, Petitpas I, Bhattacharya AA, Otagiri M, and Curry S (2005) Structural basis of the drug binding specificity of human serum albumin. *J Mol Biol* **353:**38 –52.

Gleeson MP, Waters NJ, Paine SW, and 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, and Hall SD (2013) Quantitative structure-activity relationship models of clinical pharmacokinetics: clearance and volume of distribution. *J Chem Inf Model* **50:**948–957.

Goodman LS, and Gilman A (2017) *The Pharmacological Basis of Therapeutics*, 13th ed, McGraw-Hill Publishers, New York.

Grime K, and Riley RJ (2006) The impact of in vitro binding on in vitro-in vivo extrapolations, projections of metabolic clearance and clinical drug-drug interactions. *Curr Drug Metab* **7:**251–264.

Gunturi SB, and Narayanan R (2007) In silico ADME modeling 3: computational models to predict human intestinal absorption using sphere exclusion and kNN QSAR methods. *QSAR Combinat Sci* **26:**653– 668.

Israili ZH, Dayton PG (2001) Human alpha-1-glycoprotein and its interactions with drugs. *Drug Metab Rev* **33:** 161-235.

Jolivette LJ, and 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.

Kremer JMH, Wilting J, and Janssen LHM (1988) Drug binding to human alpha-1-acid glycoprotein in health and disease. *Pharmacol Rev* **40:**1– 47.

Lewis DFV, and Dickins M (2003) Baseline lipophilicity relationships in human cytochromes P450 associated with drug metabolism. *Drug Metab Rev* **35:**1–18.

Lipinski CA, Lombardo F, Dominy BW, and Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **23:**3–25.

Lombardo F, Obach RS, Shalaeva MY, and 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, and 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, and 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, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, Harriman SP, Berellini G, Rajlic I, and Obach RS (2013a) 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, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, Harriman SP, Berellini G, Rajlic I, and Obach RS (2013b) Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 2: clearance. *J Clin Pharmacol* **53:** 178-191.

Lombardo F, Obach RS, Varma MV, Stringer R, and Berellini G (2014) Clearance mechanism assignment and total clearance prediction in human based upon in silico models. *J Med Chem* **57:**4397–4405.

Lombardo F, and 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, Desai PV, Arimoto R, Desino KE, Fischer H, Keefer CE, Petersson C, Susanne Winiwarter S, and Fabio Broccatelli F (2017) In silico ADME-PK utility and best practices – an industry perspective from the international consortium for innovation through quality in pharmaceutical development. *J Med Chem* **60:**9097–9113.

Mahmood I, Martinez M, and Hunter RP (2006) Interspecies allometric scaling: I. prediction of clearance in large animals. *J Vet Pharmacol Ther* **29:**415– 423.

Martinez M, Mahmood I, and Hunter RP (2006) Interspecies allometric scaling: prediction of clearance in large animal species: II. Mathematical considerations. *J Vet Pharmacol Ther* **29:**425– 432.

McGinnity DF, Collington J, Austin RP, and Riley RJ (2007) Evaluation of human pharmacokinetics, therapeutic dose and exposure predictions using marketed oral drugs. *Curr Drug Metab* **8:**463– 479.

Meijer DKF, Van der Sluijs P (1987) The influence of binding to albumin and α_1-acid glycoprotein on the clearance of drugs by the liver. *Pharm Weekbl Sci* **9:** 65-74

Nestorov I, Gueorguieva I, Jones HM, Houston B, and Rowland M (2002) Incorporating measures of variability and uncertainty into the prediction of in vivo hepatic clearancefrom in vitro data. *Drug Metab Dispos* **30:**276 –282.

Norinder U, and Bergström CAS (2007) Prediction of ADMET properties. *Chem Biol* **3:**1003-1042.

Obach RS, Baxter JG, Liston TE, Silber BM, Jones BC, MacIntyre F, Rance DJ, and 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, and 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, and Tozer TN (1979) Effect of altered plasma protein binding on apparent volume of distribution. *J Pharm Sci* **68:**1203–1205.

Riley RJ, McGinnity DF, and Austin RP (2005) A unified model for predicting human hepatic, metabolic clearance from in vitro intrinsic clearance data in hepatocytes and microsomes. *Drug Metab Dispos* **33:**1304–1311.

Smith DA (1997) Physicochemical properties in drug metabolism and pharmacokinetics, in *Computer-Assisted Lead Finding and Optimisation* (van de Waterbeemd H, Testa B, Folkers G eds) pp 267–276, Wiley-VCH, Weinheim, Germany.

Stepan AF, Karki K, McDonald WS, Dorff PH, Dutra JK, DiRico KJ, Won A, Subramanyam C, Efremov IV, O'Donnell CJ, Nolan CE, Becker SL, Pustilnik LR, Sneed B, Sun H, Lu Y, Robshaw AE, Riddell D, O'Sullivan TJ, Sibley E, Capetta S, Atchison K, Hallgren AJ, Miller E, Wood A, and Obach RS (2011) Metabolism-directed design of oxetane-containing arylsulfonamide derivatives as γ-secretase inhibitors. *J Med Chem* **54:** 7772-7783.

Tang H, and Mayersohn M (2006) A global examination of allometric scaling for predicting human drug clearance and the prediction of large vertical allometry. *J Pharm Sci* **95:**1783–1799.

van de Waterbeemd H, Smith DA, and Jones BC (2001) Lipophilicity in PK design: methyl, ethyl, futile. *J Comp Aided Mol Design* **15:**273–286.

Veber DF, Johnson SR, Cheng H, Smith BR, Ward KW, and Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* **45:**2615–2623.

Ward KW and Smith BR (2004a) A comprehensive quantitative and qualitative evaluation of extrapolation of intravenous pharmacokinetic parameters from rat, dog, and monkey to humans: I. Clearance. *Drug Metab Dispos* **32:**603–611.

Ward KW and Smith BR (2004b) 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 and Lombardo F. (2010) Use of the Øie-Tozer model in understanding mechanisms and determinants of drug distribution. *Drug Metab Dispos* **38:**1159– 1165.

FIGURE LEGENDS

FIGURE 1. Distribution of human pharmacokinetic values for the 1352 compounds included in this analysis. Panel A: VD_{ss}; Panel B: Clearance; Panel C: $t_{1/2}$; Panel D: f_u.

FIGURE 2. Distribution of computed physicochemical properties for the 1352 compounds included in this analysis. Panel A: MW; Panel B: Number of rotatable bonds (NRB); Panel C: Polar Surface Area (PSA, N and O atoms only); Panel D: ionization state; Panel E: logP; Panel F: $logD_{7.4}$; Panel G: Number of hydrogen-bond acceptors; Panel H: Number of hydrogen-bond donors.

FIGURE 3. Relationship between median VD_{ss} values and computed physicochemical parameters. The median value is indicated by the horizontal line in the gray box and the lower and upper limits of the box represent the first and third quartile, respectively. The black points represent the compounds. Panel A: VD_{ss} vs. ionization state; Panel B: VD_{ss} vs. $logD_{7.4}$; Panel C: VD_{ss} vs. PSA; Panel D: VD_{ss} vs. NRB; Panel E: VD_{ss} vs. number of hydrogen-bond acceptors; and Panel F: VD_{ss} vs. number of hydrogen-bond donors.

FIGURE 4. Relationship between median clearance values and computed physicochemical parameters. The median value is indicated by the horizontal line in the gray box and the lower and upper limits of the box represent the first and third quartile, respectively. The black points represent the compounds. Panel A: clearance vs. ionization state; Panel B: clearance vs. $logD_{7.4}$; Panel C: clearance vs. PSA; Panel D: clearance vs. NRB; Panel E: clearance vs. number of hydrogen-bond acceptors; and Panel F: clearance vs. number of hydrogen-bond donors.

FIGURE 5. Relationship between median fraction unbound in plasma (f_u) values and computed physicochemical parameters. The median value is indicated by the horizontal line in the gray box and the lower and upper limits of the box represent the first and third

quartile, respectively. The black points represent the compounds. Panel A: f_u vs. ionization state; Panel B: f_u vs. $logD_{7.4}$; Panel C: f_u vs. PSA; Panel D: f_u vs. NRB; Panel E: f_u vs. number of hydrogen-bond acceptors; and Panel F: f_u vs. number of hydrogen-bond donors.

FIGURE 6. Relationship between total and free VD_{ss} or total and free clearance values and computed $logD_{7.4}$. Acidic compounds are represented by red dots, basic compounds by blue dots, neutral compounds by green dots and zwitterionic compounds by yellow dots in each panel. Panel A: Total VD_{ss}; Panel B: Free VD_{ss}; Panel C: total clearance; and Panel D: free clearance.

FIGURE 7. Distribution of total VD_{ss} and clearance values utilizing value ranges reported in Table 1 against vertical bins based on year of first disclosure. The bins after 1960 span approximately two decades and the colors indicate property range values in ascending order from blue to red or brown (bottom to top).

FIGURE 8. Trends in the distribution of f_u and computed $logD_{7.4}$ values utilizing the ranges reported in Table 1 for f_u and the same ranges as in Figure 2 for computed $logD_{7.4}$. The value ranges and frequency of compounds are reported against vertical bins based on year of first disclosure. The bins after 1960 span approximately two decades and the colors indicate property range values in ascending order from blue to red or brown (bottom to top).

FIGURE 9. Trend of molecular weight utilizing the median value (371 Da) of the entire data set as threshold vs. vertical bins reflecting years of first disclosure.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

TABLE 1. Characteristics of the human intravenous pharmacokinetic parameters and plasma protein binding for 1352 compounds in the dataset.

Parameter	N	%
VD_{ss} (L/kg)		
compounds less than 0.1	36	3
compounds between 0.1 and 0.7	525	40
compounds between 0.7 and 2	357	27
compounds between 2 and 10	289	22
compounds greater than 10	108	8
CL (mL/min/kg)		
compounds less than 1	215	16
compounds between 1 and 5	491	36
compounds between 5 and 15	428	32
compounds between 15 and 21	81	6
compounds greater than 21	135	10
Half-Life ($t_{½}$) (h)		
compounds less than 1	172	13
compounds between 1 and 4	466	35
compounds between 4 and 12	377	28
compounds between 12 and 24	146	11
compounds greater than 24	174	13
Fraction unbound in plasma (f_u)		
compounds lower than 0.01	111	12
compounds between 0.01 and 0.05	143	16
compounds between 0.05 and 0.1	106	12
compounds between 0.1 and 0.5	262	28
compounds greater than 0.5	298	32

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

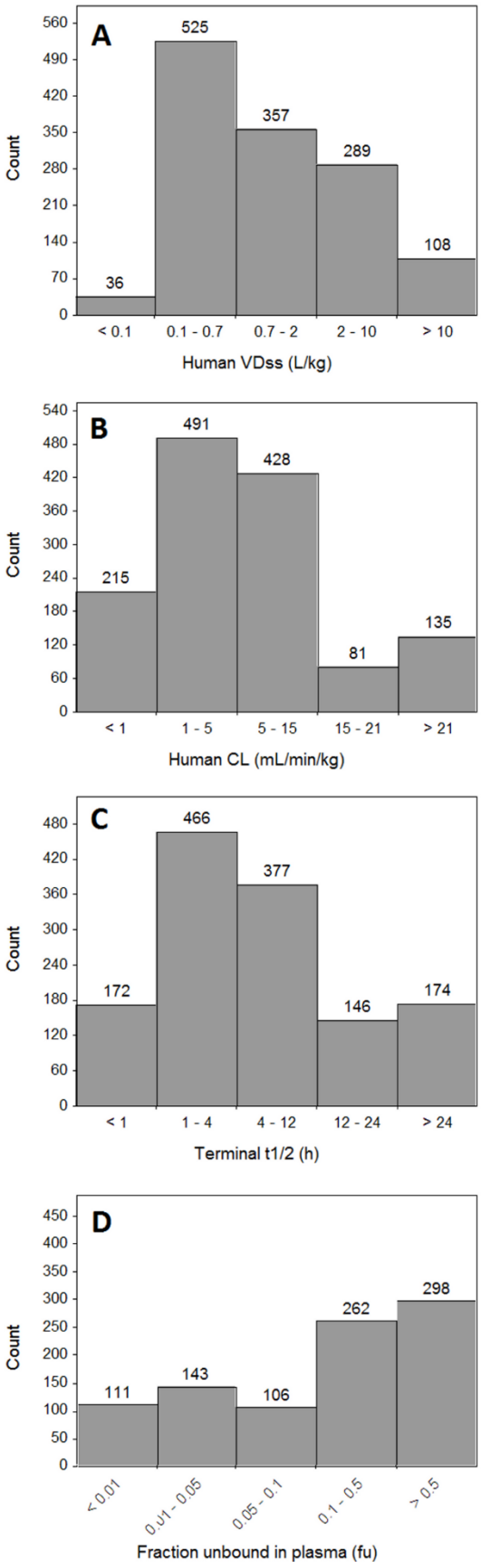

Figure 1.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

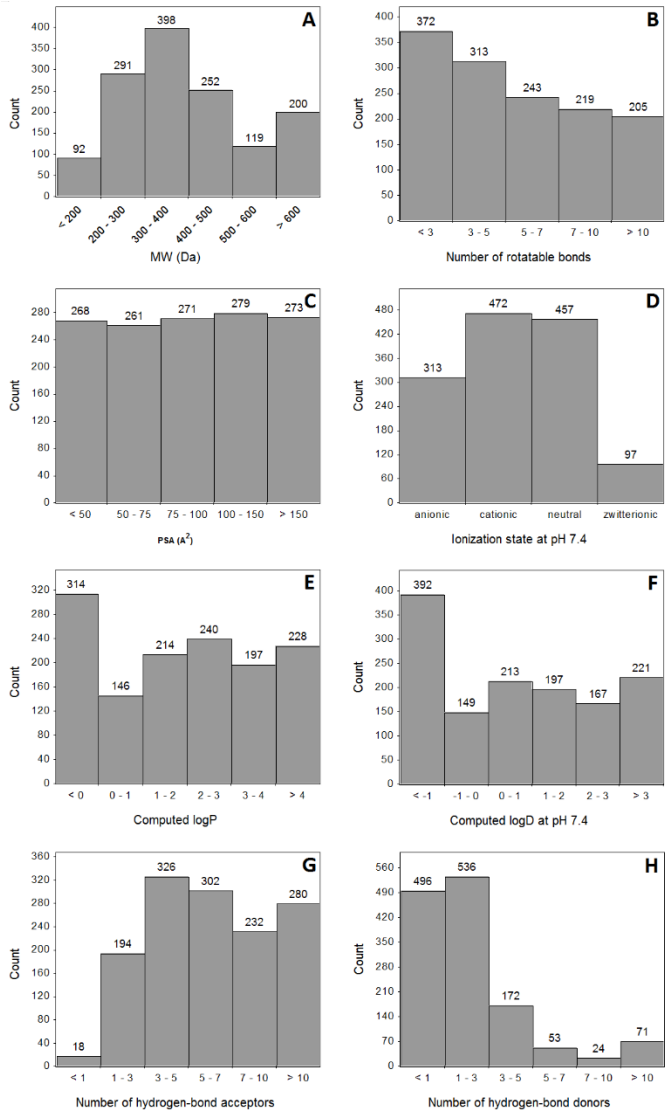

Figure 2.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

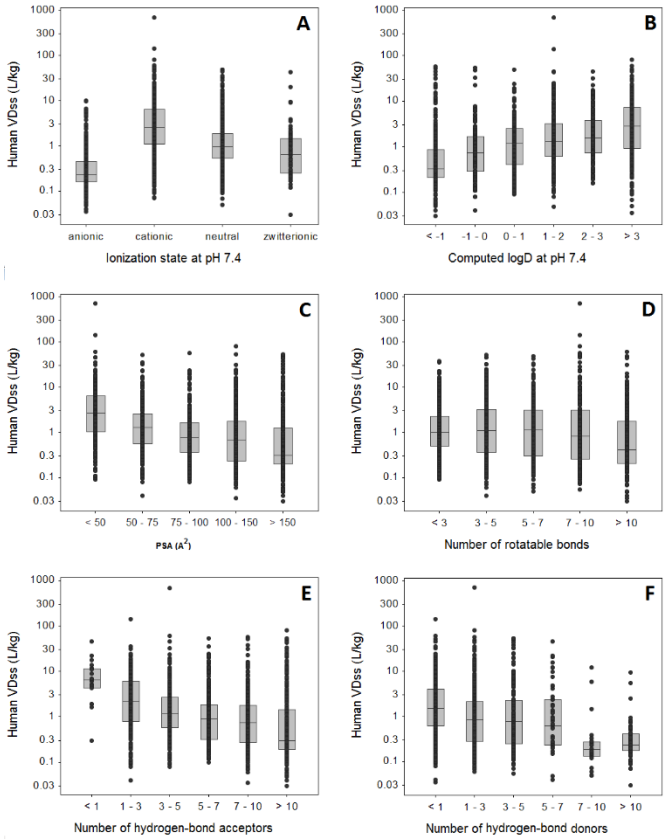

Figure 3.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

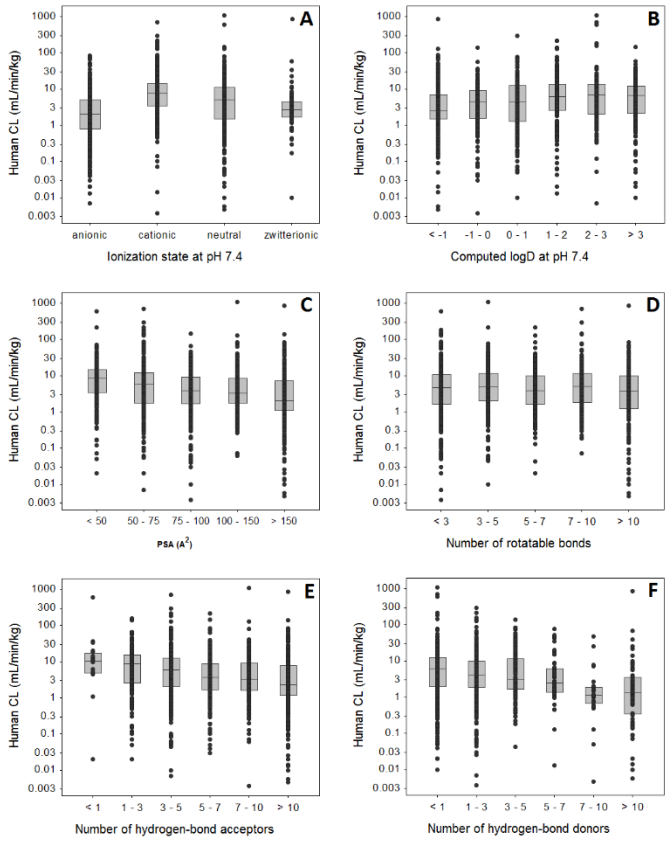

Figure 4.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

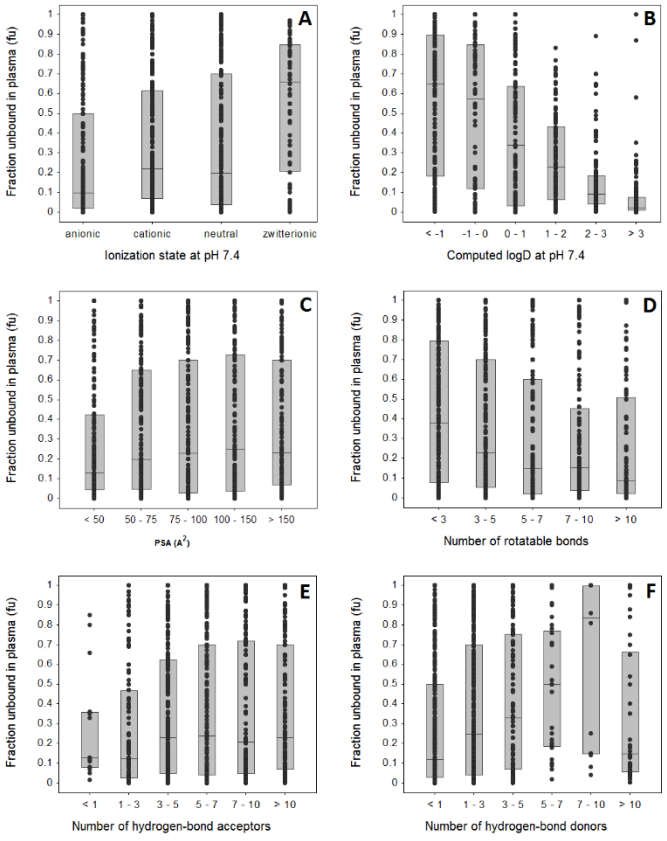

Figure 5.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

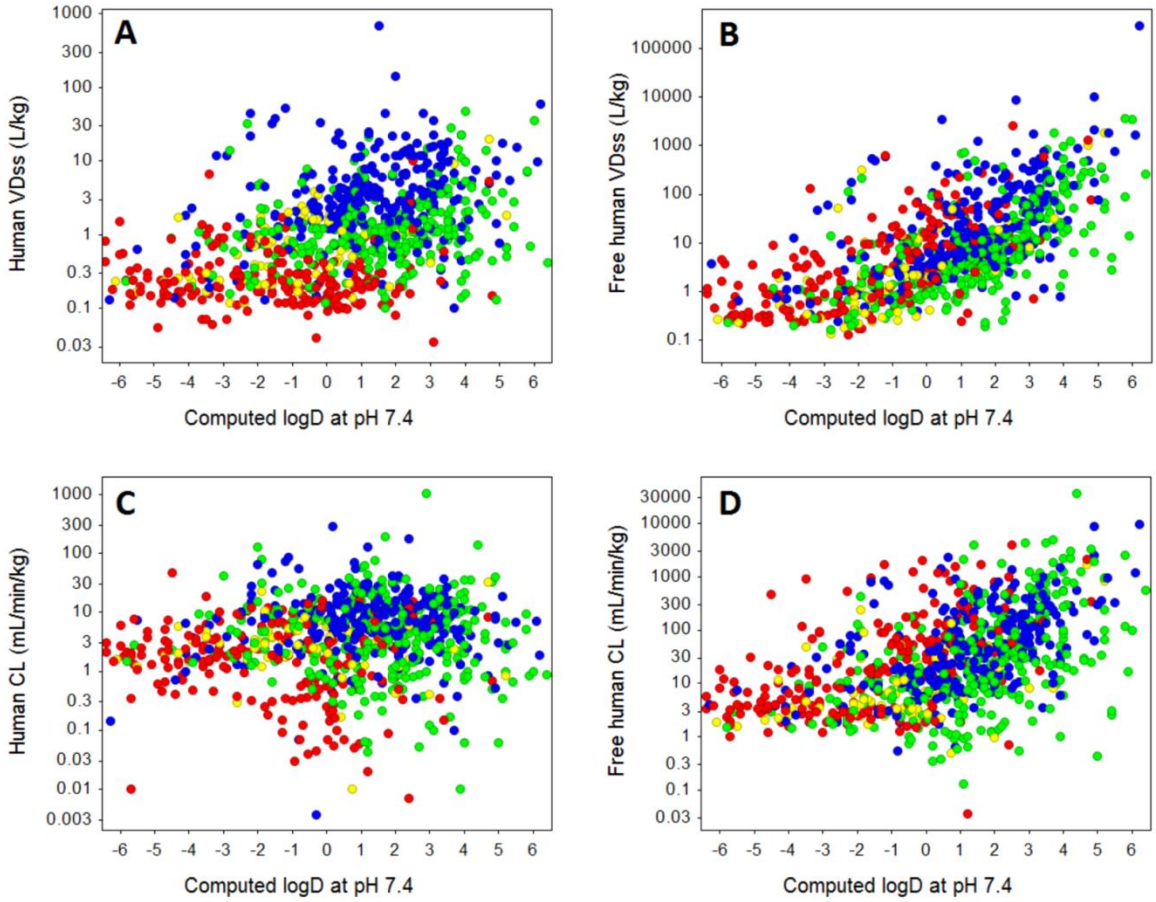

Figure 6.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

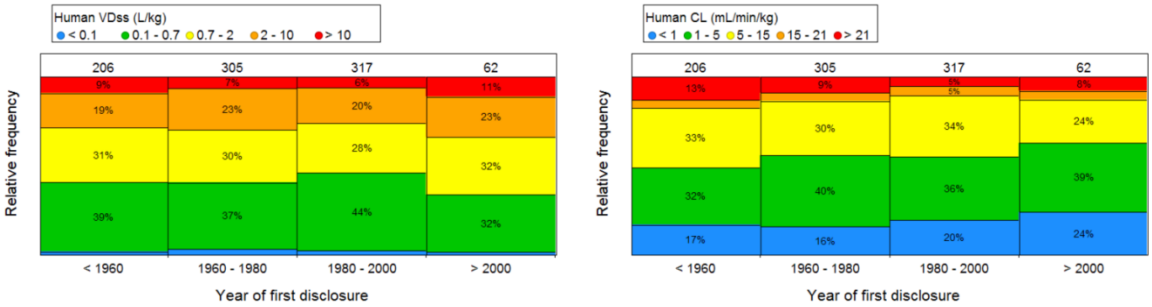

Figure 7.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

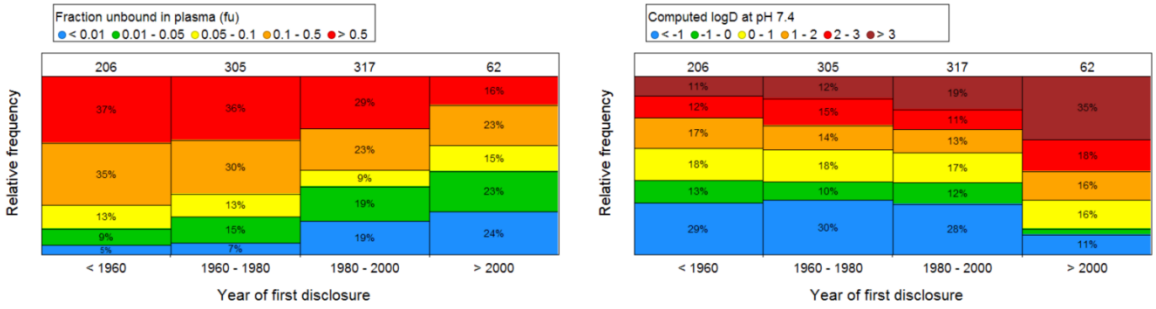

Figure 8.

DMD Fast Forward. Published on August 16, 2018 as DOI: 10.1124/dmd.118.082966
This article has not been copyedited and formatted. The final version may differ from this version.

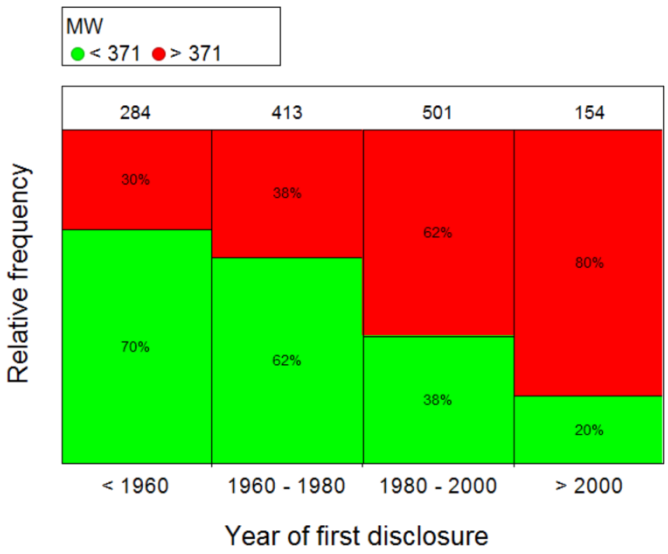

Figure 9.