Title Page

Hepatic clearance predictions from *in vitro-in vivo* extrapolation and BDDCS

Christine M. Bowman, Leslie Z. Benet

Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco (CMB, LZB)

Running Title Page

Hepatic clearance predictions and BDDCS

Corresponding author: Leslie Z. Benet, Department of Bioengineering and Therapeutic Sciences, University of California, 533 Parnassus Ave., Room U-68, San Francisco, California 94143-0912, 415-476-3853 (telephone), 415-476-8887 (fax), leslie.benet@ucsf.edu

Text pages: 24

Tables: 5

Figures: 0

References: 30

Words in abstract: 223

Words in introduction: 389

Words in discussion: 961

Abbreviations: CL_{int}, intrinsic clearance; IVIVE, *in vitro* to *in vivo* extrapolation; BDDCS, Biopharmaceutics Drug Disposition Classification System; fu_b, unbound fraction in blood; fu_{inc}, unbound fraction in incubation; AFE, average fold error; RMSE, root mean squared error; CL_{met}, metabolic clearance

Abstract

Predicting *in vivo* pharmacokinetic parameters such as clearance from *in vitro* data is a crucial part of the drug development process. There is a commonly cited trend that drugs that are highly protein bound and are substrates for hepatic uptake transporters often yield the worst predictions. Given this information, 11 different data sets using human microsomes and hepatocytes were evaluated to search for trends in accuracy, extent of protein binding, and drug classification based on the Biopharmaceutics Drug Disposition Classification System (BDDCS), which makes predictions about transporter effects. As previously reported, both *in vitro* systems (microsomes and hepatocytes) gave a large number of inaccurate results, defined as predictions falling more than 2-fold outside of *in vivo* values. The weighted average of the percentage of inaccuracy was 66.5%. BDDCS class 2 drugs, which are subject to transporter effects *in vivo* unlike class 1 compounds, had a higher percentage of inaccurate predictions and often had slightly larger bias. However, since the weighted average of the percent inaccuracy was still high in both classes (81.9% for class 2, and 62.3% for class 1), it may be currently hard to use BDDCS class to predict potential accuracy. The results of this study emphasize the need for improved IVIVE experimental methods as using physiologically based scaling is still not accurate, and BDDCS cannot currently help predict accurate results.

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

The current drug development process is expensive, time-consuming, and inefficient due to compound attrition (Pammolli et al., 2011). While failures due to pharmacokinetic parameters have decreased in recent years (Waring et al., 2015), continued improvement in pharmacokinetic predictions is crucial.

Metabolic stability studies are some of the earliest *in vitro* studies conducted during drug development to determine the rate and extent to which a molecule is metabolized, and can be useful for rank ordering candidates. After measuring *in vitro* metabolic turnover, or intrinsic clearance (CL_{int}), *in vivo* hepatic clearance can be predicted using *in vitro-in vivo* extrapolation (IVIVE) methods. A common approach is to apply physiologically based scaling factors to the raw *in vitro* data, such as hepatocellularity for studies using hepatocytes or a factor to account for incomplete microsomal recovery for microsomes, and to then apply a model of hepatic disposition such as the well-stirred model (Houston, 1994). While the results are often used in the drug development process, there is perhaps an overemphasis placed on their reliability.

The first part of the present study examined the overall accuracy of hepatic clearance predictions in the field at this time. Many groups have attempted IVIVE, tried to create new models to improve predictions from old *in vitro* values, or investigated different experimental setups. A study published 10 years ago collected and examined results from 85 compounds, concluding there was a paucity of literature data (Nagilla et al., 2006), however much work has been done since then.

When examining the accuracy of these values, a prediction bias has been found that is unresolved from human variability and experimental uncertainty (Hallifax and Houston, 2009). There is also a commonly cited trend that substrates for hepatic uptake transporters and highly

protein bound compounds yield the poorest predictions (Soars et al., 2007). The Biopharmaceutics Drug Disposition Classification System (BDDCS), which categorizes transporter effects on drug disposition, says class 1 compounds exhibit minimal clinically relevant transporter effects, while class 2 compounds may be governed by transporter effects in the gut and liver (Wu and Benet, 2005). BDDCS has become an important part of early drug discovery for predicting routes of elimination, food effects, and potential drug interactions (Wu and Benet, 2005). Given this trend, the main objective of this study was to determine if BDDCS classification could be a determinant of accurate IVIVE results.

Materials and Methods

A literature search was conducted for compounds previously described for which both *in vitro* and *in vivo* clearance data were available. Studies using human microsomes as well as human hepatocytes were considered, as both systems are routinely used in the pharmaceutical industry. The terms used as keywords to help in the search included "*in vitro-in vivo* extrapolation", "intrinsic clearance", "microsomes", "hepatocytes", or a combination of these.

All the studies considered here used the well-stirred model in their predictions, and predictions were made using physiologically based scaling factors, not empirical or regression-based factors. The data sets were examined separately, excluding re-examination of previously published data, as different experimental setups (such as the inclusion of serum in incubations) and scaling (such as the inclusion of fu_b and fu_{inc} vs. no binding terms) were used in each. Similarly, repeated drugs were not removed due to value differences among data sets. Overall evaluations were also tabulated. The data evaluated can be found in Supplementary Table 1.

The accuracy of predictions was determined based on whether or not the predictions fell within 2-fold of the true *in vivo* values, as has been a standard cutoff in previous studies (Blanchard et al., 2006; Fagerholm, 2001; Zuegge et al., 2001).

To measure bias, the average fold error (AFE) was calculated using the following equation (Obach et al., 1997):

$$AFE=10^{\frac{1}{N}\sum \log\left(\frac{observed}{predicted}\right)}$$

AFE was recorded as the whole number reciprocal if less than 1.

The precision was also calculated with the root mean squared error (RMSE) using (Sheiner and Beal, 1981):

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum(predicted-observed)^2}$$

To divide the compounds based on their BDDCS classification, two publications categorizing over 900 drugs and over 175 drugs were consulted (Benet et al., 2011; Hosey et al., 2016). Five compounds were also classified here for the first time (class 1: amobarbital, bufuralol, levoprotiline, and triprolidine; class 2: tenidap). Trends in the accuracy of predictions compared to class 1 and class 2 drugs, where metabolism is the main route of elimination, were examined. Protein binding was also considered if the values used in the prediction calculations were available, as the interplay between protein binding, transporters, and enzymes is known to be important (Benet, 2009). Drugs with high protein binding were defined as having a free fraction less than or equal to 0.05.

Results

Seven different papers were examined that fit the criteria mentioned above (Brown et al., 2007; Hallifax et al., 2010; Ito et al., 2005; McGinnity et al., 2004; Obach et al., 1999; Riley et al., 2005; Sohlenius-Sternbeck et al., 2010). Hallifax et al. (2010) compiled a large database of predictions from many of the papers also examined here, however not all drugs from the original papers were included and often different values of $CL_{in\ vivo}$ were compared, leading the same drugs to be accurately or inaccurately predicted based on the value choices. Furthermore, while it could be argued that the more recent Hallifax et al. paper provides refined values from the original papers, looking at the percentage inaccuracy and AFE both overall and for class 1 and class 2 drugs reveals that the Hallifax et al. data often actually have comparable or higher percentage inaccuracy and AFE values compared to the original papers. All papers were therefore examined to try to obtain a fuller picture of the relationship to BDDCS. Five human microsome data sets, some with multiple scaling options, were included in this evaluation for a total of 332 values, and six human hepatocyte data sets were included also for a total of 332 values. The percentage of inaccurate predictions (more than 2-fold difference) for each data set as well as the AFE and RMSE are shown in Table 1. Every data set examined has 41.0% or greater inaccuracy and AFE values are as high as 21.7. The paper by Sohlenius-Sternbeck et al. (2010) only provided individual prediction values using a regression model so further analysis could not be conducted. However, since it is the most recent paper examined, the summary statistics using the well-stirred model with protein binding that were given were still included in the table for comparison. The weighted average for the percentage of inaccurate results for microsomes is 66.8%, for hepatocytes is 66.2%, and overall is 66.5%.

The same papers and data sets were used to examine BDDCS trends. Class 1 and class 2 drugs were compiled from each set, and the inaccuracy of the predictions, AFE, and RMSE for each class were determined (Table 2). As expected, class 2 drugs have a higher percentage of inaccurate predictions than class 1 drugs in every case except one, where all predictions were inaccurate. The AFE was either slightly higher or almost identical for class 2 drugs compared to class 1 drugs. Considering a total of 305 class 1 drug values, the weighted average of the percentage of inaccurate predictions is 62.3%. For a total of 155 class 2 drug values, the weighted average of the percentage of inaccuracy is 81.9%. (The total number of class 1 and 2 drugs is less than 644 since individual drugs are not enumerated in Sohlenius-Sternbeck et al. (2010) and some unapproved proprietary compounds are included in other data sets.) For class 1 drugs, studies done in microsomes have a weighted average of 63.3% inaccuracy, while studies in hepatocytes are 66.2% inaccurate. For class 2 drugs, studies in microsomes have a weighted average of prediction inaccuracy of 85.6%, while studies in hepatocytes have a 78.4% average.

Finally, given that substrates of transporters and highly bound drugs often have the poorest clearance predictions (Soars et al., 2007), protein-binding differences were examined between the two BDDCS classes. First, the percentage of drugs with inaccurate predictions that are also highly protein bound in both classes was determined (Table 3). There are more inaccurate class 2 drugs that are highly protein bound than class 1 drugs in every case examined. The weighted average of inaccurate class 1 drugs with high protein binding is 19.8%, while the weighted average for class 2 is 67.3%. Since class 2 drugs in general are often highly protein bound (Broccatelli et al., 2012), the numbers of highly bound drugs in both classes that have inaccurate predictions were also determined (Table 4). These results agree with several other conclusions that highly protein bound compounds are often poorly predicted. Class 1 highly

protein bound drugs were inaccurately predicted 81.3% of the time, and class 2 highly bound drugs had an 85.7% average inaccuracy rate. In four data sets, highly bound class 2 drugs had a higher percentage of inaccuracy than class 1 drugs, in one data set the opposite was true, and in the last all highly bound drugs were inaccurate.

Looking at the bias between the high and low protein binding drugs in the two classes (Table 5), it is difficult to see trends between the two classes, however the bias is always higher for the high protein binding drugs, except in the case of the data from Obach et al. (1999) using fu_b and fu_{inc}, and Brown et al. (2007) where there are only two class 1 high protein binding drugs and 4 class 2 low protein binding drugs perhaps skewing the results.

Discussion

Being able to accurately predict pharmacokinetic parameters, especially clearance, early in the drug development process is a key part of lead optimization. However while some studies have claimed to find success in predicting *in vivo* clearance from *in vitro* data, others have questioned the reliability (Masimirembwa et al., 2003). Underpredicting *in vivo* clearance may result in inefficiency in the drug discovery pipeline or an ineffective therapeutic dosing regimen, while overpredicting *in vivo* clearance may lead to missed opportunities that were rejected early in the development process (Clarke et al., 2001).

The goal of this study was to compile data to examine the accuracy of the prediction methods for *in vivo* clearance and relate this accuracy to BDDCS classification. For the 11 data sets considered, there is a large percentage of inaccuracy. To have a true understanding of the accuracy of *in vitro* methods, physiologically scaled *in vitro* estimations and observed *in vivo* clearance were directly compared, since incorporating established physiological scaling factors as well as unbound fractions in the blood and possibly *in vitro* matrix should in theory, give accurate predictions. This is in contrast to some groups creating linear regression equations from reference compound data and then applying an empirical scaling factor to try to further improve predictions (Sohlenius-Sternbeck et al., 2012). The fact that 66.5% of predictions overall are inaccurate emphasizes the idea that a mechanistic understanding of this inaccuracy still needs to be determined before IVIVE predictions can be completely trusted.

BDDCS classification and protein binding were then examined to see if they could separate accurate from inaccurate results to help determine whether predictions can be trusted in the future or not. Class 1 drugs, or those that are extensively metabolized and highly soluble, appear to overwhelm transporter effects, while class 2 drugs, also extensively metabolized but

poorly soluble, can be affected by efflux transporters in the gut, and both uptake and efflux transporters in the liver (Shugarts and Benet, 2009). Given the trend that poorly predicted compounds are often substrates for transporters (Soars et al., 2007), it was expected that class 1 drugs that have no clinically relevant transporter effects would yield better predictions than class 2 drugs. The other part of the trend is that poorly predicted compounds are also often highly protein bound, which is why protein binding was considered when data were available (Ring et al., 2011). Overall, the hypothesis was that class 2 drugs would be more poorly predicted due to the fact that they are substrates for transporters, and these poorly predicted class 2 drugs would also be highly protein bound.

As expected, class 2 drugs yielded poorer predictions in every case examined; however, there was still large inaccuracy for both class 1 and class 2 drugs. Class 2 drugs also often had a higher AFE, but not different enough (or sometimes at all) to understand bias. However, AFE provides a better measure of bias than RMSE, which is highly influenced by the marked differences in CL_{int} values from study to study. For example, the values reported by Brown et al. (2007) for predicted and measured CL_{int} for propofol were 2,773 and 5,052 ml/min/kg, respectively, while for the same drug McGinnity et al. (2004) reported 283 and 24 ml/min/kg. At this point in time with the current methodology, relying on BDDCS class cannot confidently provide information about whether predictions will be accurate or not. This agrees with previous findings from Poulin et al. who found that predictivity was similar between classes for a human microsome data set of 42 drugs (Poulin et al., 2012). It is interesting to note that microsomes and hepatocytes gave similar prediction accuracies in both class 1 and class 2 drugs. A bigger difference between the two systems would have been expected for class 2 drugs where transporters play a role since necessary uptake transporters are not present in microsomes. This

again emphasizes that there are likely major missing determinants when trying to mimic the interplay between protein binding, uptake, and metabolism *in vitro*.

Poulin et al. (2012) also suggested that an approach involving determination of effective fraction unbound in plasma based on albumin-facilitated hepatic uptake of acidic/neutral drugs improved the prediction accuracy and precision for 25 high protein binding drugs. Hallifax and Houston (2012) examined this approach for 107 drugs studied in hepatocytes and microsomes also finding an increase in prediction accuracy, but no change in precision and reported that there was no evidence that prediction bias was associated with measured fraction unbound in plasma. These latter authors emphasized the need for further "mechanistic elucidation to improve prediction methodology rather than empirical correction of bias".

Lastly, protein binding was considered along with BDDCS. Given current trends, class 2 drugs with high protein binding would have been expected to yield the poorest results. There were more inaccurate class 2 drugs that had high protein binding than class 1, but this may be because class 2 drugs generally have higher protein binding than class 1 (Broccatelli et al., 2012). This, coupled to the fact that there may be a slight dependency of bias on protein binding both here and as found previously with hepatocytes by Hallifax et al. (2010), could explain some of the difference seen between the inaccuracies in class 1 and 2 drugs. However, on average, highly bound drugs in both classes had similar high percentages of inaccuracy, and there were no clear trends in the bias or precision of highly bound drugs between classes.

This study emphasizes the fact that the *in vitro* to *in vivo* extrapolation of hepatic clearance needs to be improved through a better understanding of clearance mechanisms as *in vitro* methods on their own are often not accurate, and looking at BDDCS class cannot separate out which compounds will have accurate predictions.

Authorship Contributions

Participated in research design: Benet, Bowman

Conducted experiments: Bowman

Performed data analysis: Bowman

Wrote or contributed to the writing of the manuscript: Bowman, Benet

References

Benet, LZ (2009) The drug transporter-metabolism alliance: uncovering and defining the interplay. *Mol Pharm* **6:**1631-1643.

Benet LZ, Broccatelli F, Oprea TI (2011) BDDCS applied to over 900 drugs. *AAPS J* **13:**519-547.

Blanchard N, Hewitt NJ, Silber P, Jones H, Coassolo P, Lavé T (2006) Prediction of hepatic clearance using cryopreserved human hepatocytes: a comparison of serum and serum-free incubations. *J Pharm Pharmacol* **58:**633-641.

Broccatelli F, Cruciani G, Benet LZ, and Oprea TI (2012) BDDCS class prediction for new molecular entities. *Mol Pharm* **9:**570-580.

Brown HS, Griffin M, and Houston JB (2007) Evaluation of cryopreserved human hepatocytes as an alternative in vitro system to microsomes for the prediction of metabolic clearance. *Drug Metab Dispos* **35:**293-301.

Clarke SE and Jeffrey P (2001) Utility of metabolic stability screening: comparison of in vitro and in vivo clearance. *Xenobiotica* **31:**591-598.

Fagerholm U (2007) Prediction of human pharmacokinetics—improving microsome-based predictions of hepatic metabolic clearance. *J Pharm Pharmacol* **59:**1427-1431.

Hallifax D and Houston JB (2009) Methodological uncertainty in quantitative prediction of hepatic clearance from human experimental systems. *Curr Drug Metab* **10:**307-321.

Hallifax D and Houston JB (2012) Evaluation of hepatic clearance prediction using *in vitro* data: emphasis on fraction unbound in plasma and drug ionization using a database of 107 drugs. *J Pharm Sci* **101:**2645-2652.

Hallifax D, Foster JA, and Houston JB (2010) Prediction of human metabolic clearance from *in vitro* systems: retrospective analysis and prospective view. *Pharm Res* **27:**2150-2161.

Hosey CM, Chan R, and Benet LZ (2016) BDDCS predictions, self-correcting aspects of BDDCS assignments, BDDCS assignment corrections, and classification for more than 175 additional drugs. *AAPS J* **18:**251-260.

Houston, JB (1994) Utility of *in vitro* drug metabolism data in predicting *in vivo* metabolic clearance. *Biochem Pharmacol* **47:**1469-1479.

Ito K and Houston JB (2005) Prediction of human drug clearance from *in vitro* and preclinical data using physiologically based and empirical approaches. *Pharm Res* **22:**103-112.

Masimirembwa CM, Bredberg U, and Andersson TB (2003) Metabolic stability for drug discovery and development. *Clin Pharmacokinet* **42:**515-528.

McGinnity DF, Soars MG, Urbanowicz RA, and Riley RJ (2004) Evaluation of fresh and cryopreserved hepatocytes as in vitro drug metabolism tools for the prediction of metabolic clearance. *Drug Metab Dispos* **32:**1247-1253.

Nagilla R, Frank KA, Jolivette LJ, and Ward KW (2006) Investigation of the utility of published in vitro intrinsic clearance data for the prediction of in vivo clearance. *J Pharmacol Toxicol Methods* **53:**106-116.

Obach RS (1999) Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes *Drug Metab Dispos* **27:**1350-1359.

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.

Pammolli F, Magazzini L, and Riccaboni, M (2011) The productivity crisis in pharmaceutical R&D. *Nat Rev Drug Discov* **10:**428-438.

Poulin P, Hop CE, Ho Q, Halladay JS, Haddad S, Kenny JR (2012) Comparative assessment of *in vitro-in vivo* extrapolation methods used for predicting hepatic metabolic clearance of drugs. *J Pharm Sci* **101:**4308-4326.

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.

Ring BJ, Chien, JY, Adkison KK, Jones HM, Rowland M, Jones RD, Yates, JWT, Ku MS, Gibson CR, He H, Vuppugalla R, Marathe P, Fischer V, Dutta S, Sinha VK, Björnsson T, Lavé T, Poulin P (2011) PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 3: comparative assessment of prediction methods of human clearance. *J Pharm Sci* **100:**4090-4110.

Sheiner LB and Beal SL (1981) Some suggestions for measuring predictive performance. *J Pharmacokinet Biopharm* **9:**503-512.

Shugarts S and Benet LZ (2009) The role of transporters in the pharmacokinetics of orally administered drugs. *Pharm Res* **26:**2039-2054.

Soars MG, McGinnity DF, Grime K, and Riley RJ (2007) The pivotal role of hepatocytes in drug discovery. *Chem Biol Interact* **168:**2-15.

Sohlenius-Sternbeck AK, Afzelius L, Prusis P, Neelissen J, Hoogstraate J, Johansson J, Floby E, Bengtsson A, Gissberg O, Sternbeck J, and Petersson C (2010) Evaluation of the human prediction of clearance from hepatocyte and microsome intrinsic clearance for 52 drug compounds. *Xenobiotica* **40:**637-649.

Sohlenius-Sternbeck AK, Jones C, Ferguson D, Middleton BJ, Projean D, Floby E, Bylund J, and Afzelius L (2012) Practical use of the regression offset approach for the prediction of *in vivo* intrinsic clearance from hepatocytes. *Xenobiotica* **42:**841-853.

Waring MJ, Arrowsmith J, Leach AR, Leeson PD, Mandrell S, Owen RM, Pairaudeau G, Pennie WD, Pickett SD, Wang G, Wallace O, and Weir A (2015) An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov* **14:**475-486.

Wu C-Y and Benet LZ (2005) Predicting drug disposition via application of BCS: transport/absorption/elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* **22:**11-23.

Zuegge J, Schneider G, Coassolo P, Lavé T (2001) Prediction of hepatic metabolic clearance: comparison and assessment of prediction models. *Clin Pharmacokinet* **40:**553-63.

Footnotes

Christine Bowman was supported by the National Science Foundation Graduate Research Fellowship Program [Grant No. 1144247].

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Table 1: Percentage inaccuracy, AFE, RMSE of IVIVE predictions for 11 data sets

Author	System	# Compounds Evaluated	# Inaccurate Predictions (%)	AFE	RMSE
Brown et al. (2007)	hepatocytes	37	26 (70.3%)	4.5	6460.2
Hallifax et al. (2010)	microsomes	68	53 (77.9%)	5.2	3708.6
	hepatocytes	89	60 (67.4%)	3.9	3137.7
Ito et al. (2005)	microsomes	52	45 (86.5%)	7.9	1337.0
McGinnity et al.[a] (2004)	hepatocytes	44	22 (50.0%)	1.4	94.1
Obach et al. (1999)	microsomes (fu_b and fu_{inc})	29	13 (44.8%)	2.3	4.9
	microsomes (fu_b)	29	22 (75.9%)	4.3	6.8
	microsomes (no binding)	29	13 (44.8%)	1.5	4.3
Riley et al.[b] (2005)	microsomes	37	27 (73.0%)	3.3	2314.2
	hepatocytes	56	38 (67.9%)	3.1	1356.5
	hepatocytes (w/ serum)	14	14 (100.0%)	21.7	2124.3
Sohlenius-Sternbeck et al.[c] (2010)	microsomes (fu_b and fu_{inc})	44	70.0%	3.8	5.8
	hepatocytes (fu_b and fu_{inc})	46	89.0%	5.9	8.0
	microsomes (no binding)	44	41.0%	0.5	6.1
	hepatocytes (no binding)	46	41.0%	0.8	5.4

[a]=CL_{int} data were evaluated

[b]=$CL_{int, ub, in\ vivo}$ data were evaluated

[c]=individual values for predictions with well-stirred model were not presented, only summary statistics

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Table 2: Percentage inaccuracy, AFE, RMSE of IVIVE predictions for BDDCS class 1 and class 2 drugs

Author	System	# Class 1 Drugs	# Inaccurate Class 1 Predictions (%)	AFE	RMSE	# Class 2 Drugs	# Inaccurate Class 2 Predictions (%)	AFE	RMSE
Brown et al.	hepatocytes	24	14 (58.3%)	3.0	294.5	12	11 (91.7%)	7.4	11335.9
Hallifax et al.	microsomes	42	30 (71.4%)	5.2	4521.7	22	20 (91.0%)	4.7	1834.4
	hepatocytes	55	36 (65.5%)	4.0	3976.5	30	22 (73.3%)	3.7	466.1
Ito et al.	microsomes	32	27 (84.4%)	6.8	390.8	16	15 (93.8%)	11.2	2312.3
McGinnity et al.	hepatocytes	32	16 (50.0%)	1.1	99.3	9	5 (55.6%)	3.0	90.9
Obach et al.	microsomes (fu_b and fu_{inc})	20	7 (35.0%)	1.9	4.6	9	6 (66.6%)	3.2	5.4
	microsomes (fu_b)	20	13 (65.0%)	3.7	6.9	9	9 (100.0%)	6.0	6.7
	microsomes (no binding)	20	7 (35.0%)	1.2	4.0	9	6 (66.7%)	2.5	4.8
Riley et al.	microsomes	24	16 (66.7%)	2.7	2399.1	11	9 (81.8%)	6.0	2298.5
	hepatocytes	28	16 (57.1%)	2.4	175.7	22	18 (81.8%)	3.8	2125.8
	hepatocytes (serum)	8	8 (100.0%)	9.6	251.0	6	6 (100.0%)	64.2	3232.0

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Table 3: Percentage inaccuracy of BDDCS class 1 and class 2 drugs that are highly protein bound

Author	System	# Inaccurate Class 1 Predictions	# Inaccurate Highly Protein Bound Class 1 Predictions (%)	# Inaccurate Class 2 Predictions	# Inaccurate Highly Protein Bound Class 2 Predictions (%)
Brown et al.	hepatocytes	14	1 (7.1%)	11	7 (63.6%)
Hallifax et al.	microsomes	30	6 (20.0%)	20	9 (45.0%)
	hepatocytes	36	9 (25.0%)	22	15 (68.2%)
Obach et al.	microsomes (fu_b and fu_{inc})	7	1 (14.3%)	6	4 (66.6%)
	microsomes (fu_b)	13	1 (7.7%)	9	4 (44.4%)
	microsomes (no binding)	7	2 (28.6%)	6	4 (66.7%)
Riley et al.	hepatocytes	16	4 (25.0%)	18	17 (94.4%)
	hepatocytes (serum)	8	2 (25.0%)	6	6 (100.0%)

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Table 4: Percentage of highly protein bound BDDCS class 1 and class 2 drugs that are inaccurate

Author	System	# Highly Protein Bound Class 1 Drugs	# Inaccurate Highly Protein Bound Class 1 Predictions (%)	# Highly Protein Bound Class 2 Drugs	# Inaccurate Highly Protein Bound Class 2 Predictions (%)
Brown et al.	hepatocytes	2	1 (50.0%)	8	7 (87.5%)
Hallifax et al.	microsomes	8	6 (75.0%)	10	9 (90.0%)
	hepatocytes	9	9 (100.0%)	20	15 (75.0%)
Obach et al.	microsomes (fu_b and fu_{inc})	2	1 (50.0%)	4	4 (100.0%)
	microsomes (fu_b)	2	1 (50.0%)	4	4 (100.0%)
	microsomes (no binding)	2	2 (100.0%)	4	4 (100.0%)
Riley et al.	hepatocytes	5	4 (80.0%)	21	17 (81.0%)
	hepatocytes (serum)	2	2 (100.0%)	6	6 (100.0%)

DMD Fast Forward. Published on August 12, 2016 as DOI: 10.1124/dmd.116.071514
This article has not been copyedited and formatted. The final version may differ from this version.

Table 5: AFE and RMSE of high and low protein binding BDDCS class 1 and class 2 drugs

Author	System	Protein Binding	Class 1		Class 2	
			AFE	RMSE	AFE	RMSE
Brown et al.	hepatocytes	high	2.0	56.4	6.3	13882.7
		low	3.1	307.1	10.3	229.6
Hallifax et al.	microsomes	high	7.8	10335.3	5.3	2671.0
		low	4.8	349.9	4.2	473.3
	hepatocytes	high	12.1	9814.8	4.2	479.9
		low	3.3	242.7	2.9	437.0
Obach et al.	microsomes (fu_b and fu_{inc})	high	1.7	0.3	4.7	3.1
		low	2.0	4.9	2.3	6.7
	microsomes (fu_b)	high	4.7	0.4	7.3	3.1
		low	3.6	7.3	5.2	8.6
	microsomes (no binding)	high	13.7	1.5	7.7	6.8
		low	1.12	17.7	1.0	2.2
Riley et al.	hepatocytes	high	3.1	175.2	3.9	2175.6
		low	2.2	173.3	2.8	136.5
	hepatocytes (serum)	high	17.0	406.3	64.2	3232.0
		low	8.0	170.2	-	-