Utility of Structural Information to Predict Drug Clearance from *in Vitro* Data

Soyoung Lee¹ and Dongsup Kim¹*

¹Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, S. Korea

Subject areas: Cheminformatics, Chemical biology

Author contribution: Soyoun Lee: main idea, experimenter. Dongsup Kim: discussion and comment.

*Correspondence and requests for materials should be addressed to D.S.K. (kds@kaist.ac.kr).

Reviewer: Kyung Tai No, Yonsei University, Republic of Korea. Hyun Joo, Inje Univ. College of Medicine, Republic of Korea.

Editor: Sun Choi, Ewha Womans University, Republic of Korea.

Received Feb 18, 2010; Revised May 20, 2010; Accepted May 30, 2010; Published June 01, 2010


Competing interests: All authors declare no financial or personal conflict that could inappropriately bias their experiments or writing.

Copyright: This article is licensed under a Creative Commons Attribution License, which freely allows to download, reuse, reprint, modify, distribute, and/or copy articles as long as a proper citation is given to the original authors and sources.

**SYNOPSIS**

In the present research, we assessed the utility of the structural information of drugs for predicting human *in vivo* intrinsic clearance from *in vitro* intrinsic clearance data obtained by human hepatic microsome experiment. To compare with the observed intrinsic clearance, human intrinsic clearance values for 51 drugs were estimated by the classical methods using *in vivo-in vitro* scale-up and by the new methods using the *in vitro* experimental data and selected molecular descriptors of drugs by the forward selection technique together. The results showed that taking consideration of molecular descriptors into prediction from *in vitro* experimental data could improve the prediction accuracy. The *in vitro* experiment is very useful when the data can estimate *in vivo* data accurately since it can reduce the cost of drug development. Improvement of prediction accuracy in the present approach can enhance the utility of *in vitro* data.

**Keywords:** prediction, intrinsic clearance, *in vitro* data, molecular descriptors, forward selection, multiple linear regression
Introduction

Knowledge of pharmacokinetic parameters provides key information for screening drug candidates at the early stage of drug discovery process to reduce the risk of late-stage attrition (van de Waterbeemd & Gifford, 2003). Intrinsic clearance, the total ability of the liver to metabolize a drug in the absence of flow limitations, is one of the most meaningful pharmacokinetic parameters since it is a distinct characteristic of a particular drug, reflecting the inherent ability of the liver to metabolize the drug (Shar & Yu, 1993).

There were a number of attempts to predict in vivo hepatic clearance and intrinsic clearance from in vitro data mainly using multivariate statistics (Schneider et al., 1999) or in vivo-in vitro correlation (Houston, 1994; Ito & Houston, 2004; Ito & Houston, 2005). They are useful and simple, yet their prediction accuracies are not sufficiently high. Recently, Ito and Houston (Ito & Houston, 2004) reported that the performance of prediction using empirical scaling factor to search for the correlation of in vitro data and in vivo drug clearance was better than using physiologically based scaling factor and allometric approaches. In the present approach, we compared the prediction accuracies of new methods to that of methods using empirical scaling factor and physiologically based scaling factor. In a earlier approach, Schneider et al. (Schneider et al., 1999) achieved satisfactory predictions by combining of in vitro clearances on human and rat experimental data using a few statistical methods, multiple linear regression, partial least squares regression, and artificial neural networks. The information of not only human in vitro data but also rat in vitro data contributed to improvement of prediction accuracy. In the present approach, the information of molecular descriptors replaced additional information, rat in vitro data, in their method. Molecular descriptors are various numerical values that characterize properties of molecules by manipulation of chemical structural information. These values was applied to predict various pharmacokinetic parameters (Clark, 2003; Didiapietris et al., 2003; Doniger et al., 2002; Hou & Xu, 2003; Langowski & Long, 2002; Lewis & Dickins, 2002; Lipinski et al., 2001; Norinder & Haeblerlein, 2002; Sugawara et al., 1998) and it was demonstrated that when these descriptors were used with animal in vivo data to predict human hepatic clearance, the prediction accuracies of the models were improved (Jolivet & Ward, 2005; Nagilla & Ward, 2004; Wajima et al., 2002). Moreover, it was reported that the information of molecular descriptors improved the prediction accuracy of human hepatic clearance from in vitro data obtained by human microsome and hepatocyte experiments (Lee & Kim, 2007) recently.

The aim of the current research is to validate the utility of the structural information of drugs for predicting intrinsic clearance from in vivo data obtained by the experiments using human microsome. Human intrinsic clearance values were estimated by the previous methods using in vivo-in vitro correlation and by the present methods using in vitro experimental data and the information of molecular descriptors of drugs together, and then compared with the observed values. By presenting better performance of the present approach, we suggest to use molecular descriptors of drugs for predicting in vivo intrinsic clearance from in vitro data.

Materials and Methods

Data Collection

Dataset consists of 51 compounds obtained from literature (Ito & Houston, 2005) that includes in vitro human microsome experimental data and in vivo intrinsic clearance values. Among 52 compounds of the original dataset in literature, the YW796 whose molecular descriptors could not be obtained were excluded in the dataset.

Prediction Models

Human in vivo intrinsic clearance values (CLint, human, in vivo) were predicted by using only in vitro values (CLint, human, in vivo) and by the new method using molecular descriptors of drug together with in vitro data. A previous prediction model, method A uses human in vitro data and an in vitro-in vivo scaling factor(Ito & Houston, 2005), which is determined by the linear regression analysis. The new prediction model, method B uses both in vitro intrinsic clearance data and molecular descriptors of drug as the explanatory variables of multiple linear regression analysis. Molecular descriptors were calculated by preADME software. Useful variables among 1078 calculated molecular descriptors were selected by forward selection technique based on the r-squared (R2) to build multiple linear regression. That is, the selection procedure that chooses the variable of the highest R2 was repeated until there was no remaining variable that increased the R2 value of the regression model. By this process, 15 descriptors were selected, and they were listed in Table 1. Method C uses only selected molecular descriptors by forward selection without in vitro data. This model was generated the importance of in vitro clearance to predict in vivo drug clearance. Method C selected 10 descriptors as listed in Table 2.

Accuracy of predictions

The squared correlation coefficient (r²) was obtained to observe the correlation between observed human in vivo intrinsic clearance and the predicted value(Hayter, 2002). For measuring prediction error, mean square error (MSE) was compared. Complete leave-one-out procedures were performed to assess the generalization ability of the suggested models. In this work, the number of samples was small, and thus we measured prediction accuracy by complete leave-one-out instead of splitting samples into train and testing drugs.

| Table 1. Selected molecular descriptors in method B |
| --- | --- |
| order | Molecular descriptor |
| 1 | ATS_Geary_08_electronegativity |
| 2 | ATS_Moreau_Bruto_06_mass |
| 3 | Graph_radius |
| 4 | No_N_oxide_groups |
| 5 | ATS_Moran_01_mass |
| 6 | AlogP98_026_C |
| 7 | ATS_Geary_10_mass |
| 8 | ATS_Moran_07_E_state |
| 9 | ATS_Geary_02_polarizability |
| 10 | Quadratic_index |
| 11 | E_state_SddC |
| 12 | ATS_Geary_04_E_state |
| 13 | ATS_Moreau_Bruto_00_polarizability |
| 14 | AI_SdsssP |
| 15 | AlogP98_019_C |
The utility of molecular descriptors for the prediction
accuracies of in vivo data does not indicate that these descriptors are strongly correlated with human acute toxicity. To examine whether they are correlated with acute toxicity or whether they can explain the difference between in vitro and in vivo values, we compared the prediction accuracies of methods B and C. In contrary to method B, method C uses only molecular descriptors without any experimental data to build the model. In Table 3, method C showed considerable accuracy to method A; however, method B showed much higher accuracy than method C. Therefore, it is reasonable that well-selected combination of molecular descriptors could provide information on human intrinsic clearance, and they are more useful to improve the utility of in vitro experimental data.

Molecular descriptors have been employed to predict various human pharmacokinetic parameters such as oral absorption, bioavailability, brain penetration, clearance, volume of distribution, and metabolism by cytochrome P450 (Clark, 2003; Didziapetris et al., 2003; Doniger et al., 2002; Hou & Xu, 2003; Langowski & Long, 2002; Lewis & Dickens, 2002; Lipinski et al., 2001; Norinder & Haebерlein, 2002; Sugawara et al., 1998). Many researches have achieved sufficient prediction accuracy using them. It has been shown that when these descriptors were used with animal in vivo data to predict the human hepatic clearance, the prediction accuracies of the models were markedly improved (Jolivette & Ward, 2005; Nagilla & Ward, 2004; Wajima et al., 2002). Similarly, considering the information of molecular descriptors together with in vitro intrinsic clearance achieved improvement of prediction performance in the present research. This improvement could be made by correcting the difference between in vivo value and in vitro data originated from pharmacokinetic factors in human body.

The extension of this work

In this work, we used the simple constitutional and molecular descriptors such as molecular weight, the number of aromatic rings, 2-D surface, and 2-D volume, and the physicochemical properties such as hydrophobicity. More extensive descriptors based on 3-D structures of chemicals could improve prediction accuracy.

This work used human microsome experimental data. The in vitro clearance can be measured by microsome and hepatocyte experiments. They show a little difference because of the differences in the free fraction (Chuang et al., 2006). Using both two experiments together with molecular descriptors could help to improve the utility of in vitro experiments.

In addition, this algorithm could be used to predict other important

Results & Discussion

Linear correlation between the log values of in vitro intrinsic clearance with the log values of in vivo intrinsic clearance

The basic assumption in all current prediction models is that the log values of in vitro intrinsic clearance are linearly correlated with the log values of in vivo intrinsic clearance based on the linear tendency in Figure 1. Prediction method B that also uses this assumption resulted in reasonably accurate prediction. We applied the multiple linear regression analysis to find out the linear correlation of the log values of in vitro intrinsic clearance and predict in vivo intrinsic clearance.

Prediction accuracy of each method

Complete leave-one-out procedures were performed to assess the generalization ability of the previous and suggested models. Prediction accuracy of each method are displayed in Table 3. Method A using only human in vitro data showed insufficient accuracy, nevertheless suggested model (method B) had better performance with significant improvement in prediction accuracy. This result is consistent with our previous research (Lee & Kim, 2007). Both human hepatic clearance and human intrinsic clearance, molecular descriptors could improve prediction accuracies of in vivo values from in vitro data.

The utility of molecular descriptors for the prediction

Our next question is the major role of molecular descriptors to improve the performance. Molecular descriptors characterize the properties of molecules (Leach & Gillet, 2003). The accuracy
pharmacokinetic parameters using in vitro experiments.

Conclusion

Intrinsic clearance is one of the most important pharmacokinetic parameters since it provides the information about liver metabolizing enzymes that is a main factor to determine the dose of orally administered drugs when considering the first-pass effect.

In the present approach, new explanatory variables for predicting human in vivo intrinsic clearance, the information of molecular descriptors, were employed to improve accuracy. Lower prediction error and stronger correlation than the previous method using simple scale-up indicated that the information of molecular descriptors helps to increase prediction accuracy of human in vivo intrinsic clearance from human in vitro intrinsic clearance data. Such improvement might be made by correcting the difference between in vivo value and in vitro data originated from human pharmacokinetic factors.

Several important human pharmacokinetic parameters such as hepatic clearance and acute toxicity cannot be obtained easily by the ethical problem of human test. Therefore, predictions and in vitro tests are very useful to predict those parameters. The present research is very useful in that it can predict human pharmacokinetic parameters accurately using computations and in vitro experimental data, and eventually, it can improve the utility of in vitro experiments. We suggest that the present approach can be applied to the prediction of any other human pharmacokinetic parameters via in vitro experimental data.

References