J Comput Aided Mol Des (2007) 21:665–673
DOI 10.1007/s10822-007-9124-0

Development of in silico models for human liver microsomal stability

Pil H. Lee · Lourdes Cucurull-Sanchez · Jing Lu · Yuhua J. Du

Received: 13 March 2007 / Accepted: 1 June 2007 / Published online: 29 June 2007

Abstract We developed highly predictive classification models for human liver microsomal (HLM) stability using the apparent *intrinsic* clearance ($CL_{int,\ app}$) as the end point. HLM stability has been shown to be an important factor related to the metabolic clearance of a compound. Robust in silico models that predict metabolic clearance are very useful in early drug discovery stages to optimize the compound structure and to select promising leads to avoid costly drug development failures in later stages. Using Random Forest and Bayesian classification methods with MOE, E-state descriptors, ADME Keys, and ECFP_6 fingerprints, various highly predictive models were developed. The best performance of the models shows 80 and 75% prediction accuracy for the test and validation sets, respectively. A detailed analysis of results will be shown, including an assessment of the prediction confidence, the significant descriptors, and the application of these models to drug discovery projects.

Keywords Human liver microsomal stability · In silico model · Prediction confidence

P. H. Lee (✉)
Computer-Assisted Drug Discovery, Pfizer Global Research and Development, Ann Arbor, MI 48105, USA
e-mail: pil.h.lee@pfizer.com

L. Cucurull-Sanchez
Pharmacokinetics, Dynamics and Metabolism, Pfizer Ltd., Ramsgate Road, Sandwich, Kent CT139NJ, UK

J. Lu · Y. J. Du
Computational Chemistry and Molecular Property Group, Groton, CT 06340, USA

Introduction

Metabolism is the major elimination pathway for most drugs. The majority of drug metabolism takes place in the liver, since many drug metabolizing enzymes are produced there. Human liver microsomal (HLM) stability has been shown to be an important factor related to the metabolic clearance of a compound. Microsomes are vesicles of the smooth endoplasmic reticulum of hepatocyte, and the primary location for the metabolizing enzymes. The microsomal assay measures the clearance of a compound due to Phase I, including cytochrome P450 (CYP) -mediated metabolism [1].

The key issues in drug metabolism include identifying the enzyme(s) involved, the site(s) of metabolism, the resulting metabolite(s), and the rate of metabolism. Methods for predicting human drug metabolism from in vitro and computational methodologies and determining relationships between the structure and metabolic activity of molecules are also critically important for understanding potential drug–drug interactions and toxicity. There are numerous experimental and computational approaches that have been developed in order to predict human metabolism which have their own limitations [2]. Robust in silico models that predict metabolic clearance are very useful in the early drug discovery stages to optimize the compound structure and to select promising leads to avoid costly drug development failures in later stages. The models can be utilized to virtually screen a large combinatorial library of compounds to evaluate for potential metabolic liability before synthesis. Appropriate use of in silico models can be an alternative to in vitro screening alone to save time and money in drug discovery and development.

In this paper, we will describe our efforts to develop various in silico models for metabolic stability of drugs in

human liver microsomes based on the in vitro assay data. We will compare the performances of a variety of models generated using different statistical methods based on several sets of descriptors. A detailed analysis of results will be shown, including an assessment of the prediction confidence, the significant descriptors and the use of these models in drug discovery projects.

Datasets

Since it was our goal to develop robust in silico models that could be used company-wide, we compiled data from assays run at multiple laboratories within the company so that the broadest chemical spaces possible could be included.

To ensure the consistency of the apparent intrinsic clearance ($CL_{int,\ app}$) measured at each laboratory and at the different time period, we only used data that were measured using an internally harmonized in vitro assay protocol [3]. We applied an internally developed standard data collection, retrieval and cleaning procedure to ensure data integrity and accuracy. For data cleaning process, suspicious data (such as $CL_{int,\ app} < 0$), contaminated data, (such as $t_{1/2}$ populated as $CL_{int,\ app}$), and other problematic data, were eliminated. For compounds with multiple measurements, if the standard deviation was higher than 10% of the average value (i.e., higher than the experimental error), the data were discarded. The histogram of remaining measurements on 14,557 molecules is shown in Fig. 1. In the figure, the percentage of the compounds in the whole dataset is plotted against the $CL_{int,\ app}$. The plot peaks at $CL_{int,\ app}$ = 20 μl/min/mg, hinting that this value can be used as a cutoff for a binary classification model. Indeed, this value was used to classify compounds: a molecule is considered to be metabolically 'stable' if $CL_{int,\ app}$ is below 20 μl/min/mg and 'unstable' otherwise.

The dataset was partitioned into training (11,646 compounds; 80%) and test (2,911 compounds; 20%) sets with

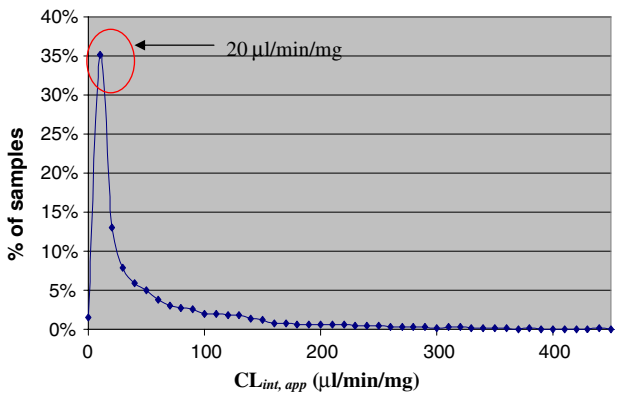

Fig. 1 Histogram of HLM data

20% (2,911 molecules) chosen by random selection using the JMP [4] software. The independent test set was used to evaluate prediction accuracies and limitations of the models developed using the training set. The training set contained 44% (5,167) stable and 56% (6,479) unstable compounds. The composition was very similar in the test set: 44% (1,282) stable and 56% (1,629) unstable compounds.

After models were built and evaluated using the training and test sets, a set of additional data was collected to validate the predictive power of the models. This validation set contained a similar distribution of stable/unstable compounds (276; 38%/450; 62%).

Modeling

Molecular descriptors

Three sets of descriptors were calculated. For the random forest analyses, a set of 2D MOE descriptors [5] as well as E-state descriptors [6] were calculated using MOE 2005.06 from the Chemical Computing Group. A set of 411 ADME Keys [7], which is a combination of a selected set of 56 MOE 2D descriptors and a set of 355 SMARTS keys were calculated using an internal program. These SMARTS keys were developed internally, and are based on variety of sources such as common functional groups, rings liable to metabolism, ISIS Keys [8], some relevant fragments and heterocyclic rings.

For the naïve Bayesian classifier studies, the Extended Connectivity Fingerprints of maximum diameter 6 (ECFP_6), as implemented in Pipeline Pilot by Scitegic [9, 10] were used. These 2D fingerprints are derived from a variant of the Morgan Algorithm [11], which indexes the environment of every atom in a molecule using up to four billion different structural features. The code used to describe the atomic features for the heavy (non-hydrogen) atoms is based on the element type, the charge, the atomic mass, and the number of connections.

Statistical methods

For statistical modeling, the Random Forest [12, 13] v4.4-2 in the R package [14] v2.0.1 and the Bayesian Classifier [15] as implemented in the Pipeline Pilot program were used.

Random forest method

Random Forest is a powerful tool for QSAR modeling, and has been successfully applied to model many biological and ADME end points. It is an ensemble of recursive partitioning trees created using bootstrap samples of the

training data set and random feature selection in tree induction. Each tree provides a classification (vote for that class). The method chooses the classification having the most votes (over all the trees in the forest). Compared to a single decision tree method, the Random Forest method generally performs better in terms of prediction accuracy. Additional nice features of the Random Forest method are its out-of-bag performance estimate that can be used in place of cross-validation, and a measure of descriptor importance that can aid in descriptor selection.

We generated a set of 500 trees with each tree giving a vote of stable or unstable for each compound. The resulting final class prediction for each molecule is whatever the majority of the trees predicted. The probability of class membership is calculated by the number of votes for that class divided by the total number of votes (500 in this case).

This probability can be used as a prediction confidence. For example, the closer the probability of being stable is to 1 the more confident the prediction is.

A naïve Bayesian classifier in combination with ECFP_6 fingerprints

The ECFP_6 fingerprints and the HLM clearance values of the training set were used to build a classification model based on naïve Bayesian statistics. A Bayesian classifier compares the frequency of occurrences of features that are found in two or more groups to find those features that discriminate best between these groups. In this study, the 'Cross-validated Learn Good Molecules (naïve Bayesian)' component in Pipeline Pilot 4.5 was used to build a classifier capable of discriminating stable compounds (i.e. with $CL_{int,\ app} < 20$ µl/min/mg) from the rest. The number of cross-validation groups was set to 5.

Calculation of statistical parameters

To evaluate the performance of the models developed, we calculated several statistical parameters as shown in Table 1. For example, sensitivity is calculated by the number of experimentally stable compounds that are predicted to be stable (A) divided by the total number of experimentally stable compounds (A + B). Specificity is the same for unstable compounds. Positive predicted value is calculated by the number of experimentally stable compounds that are predicted to be stable (A) divided by the total number of compounds predicted to be stable (A + C). Negative predicted value is calculated by the number of experimentally unstable compounds that are predicted to be unstable (D) divided by the total number of compounds predicted to be unstable (B + D). The concordance is calculated by the sum of the number of experimentally stable compounds that are predicted to be stable (A) and the number of experimentally unstable compounds that are predicted to be unstable (D) divided by the total number of compounds (N). The kappa value is calculated from the formula in the table.

Results and discussion

Three computational models developed using 2D MOE and E-state descriptors, ADME keys, and ECFP_6 fingerprints with Random Forest and Bayesian classification are compared in Table 2. In the table, the scope and limitations of three models were evaluated by deriving statistical parameters for the test and the validation sets. The concordance values show the overall prediction accuracy from each model. The Bayesian & ECFP_6 model was able to correctly classify 78%, the RF & ADME keys model 82%, and the RF & MOE, E-state model 81% of the 2,911 compounds in the test set. For the validation set, the Bayesian & ECFP_6 model, the RF & ADME keys model, and the RF & MOE, E-state model correctly classified 69, 73, and 75% of the 726 compounds, respectively. The fact that the predictions are worse for the validation set than the test set may suggest that the model should be updated continuously with new data.

For classification models, the accuracy of the models is commonly assessed by the Kappa values [16] (or Kappa statistics). The Kappa value is an index which compares the agreement against that which might be expected by chance. Kappa can be thought of as the chance-corrected proportional agreement, and possible values range from +1 (perfect agreement) via 0 (no agreement above that expected by chance) to -1 (complete disagreement). It is calculated as shown in Table 1.

The Kappa values from three models range from 0.57 to 0.64 for the test set and from 0.34 to 0.46 for the validation set. The sensitivity values indicate the percentage of stable compounds correctly classified, and the specificity values, the percentage of unstable compounds correctly classified. The sensitivity values from the three models for the test set (69–77%) are lower than the corresponding specificity values (81–88%). The same trend is observed for the validation set. This might be related to the fact that the data sets are somewhat unbalanced, with a 40/60% stable to unstable split.

The ability of the Bayesian classification model to discriminate between stable and unstable compounds was evaluated with a bimodal histogram of the test dataset, shown in Fig. 2. This histogram shows how the Bayesian scores obtained for stable compounds are distributed along the positive range, while unstable compounds tend to have negative values. It can also be seen that a grey area could

Table 1 Calculation of statistical parameters

		Prediction		
		Stable	Unstable	Total
Expt	Stable	A	B	A + B
	Unstable	C	D	C + D
	Total	A + C	B + D	N = A + B + C + D
			Concordance =	(A + D)/N
			Sensitivity =	A/(A + B)
			Specificity =	D/(C + D)
			Positive predicted value =	A/(A + C)
			Negative predicted value =	D/(B + D)
			E =	(A + C)(A + B) + (B + D)(C + D)/N × N
			O =	(A + D)/N
			Kappa =	O - E/1 - E

Table 2 Statistical parameters obtained with each modeling method for the test and validation set

Modeling method	Bayesian & ECFP_6		RF & ADME keys		RF & MOE, E-state	
Data set	Test	Validation	Test	Validation	Test	Validation
Concordance (%)	78	69	82	73	81	75
Kappa	0.57	0.34	0.64	0.43	0.61	0.46
Sensitivity (%)	75	62	77	64	75	68
Specificity (%)	81	73	86	78	86	79
Positive predictive value (%)	76	58	82	65	80	66
Negative predictive value (%)	81	76	83	78	81	80

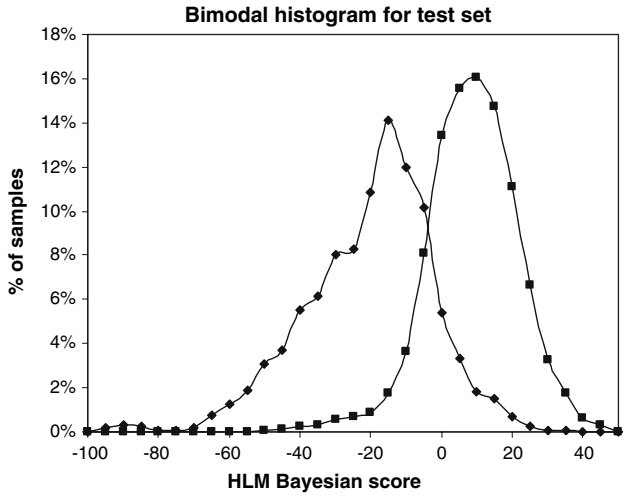

Fig. 2 Bimodal histogram representing the distribution of Bayesian scores for stable (*squares*) and unstable (*diamonds*) compounds in the test set

be defined between -10 and 0, where there is some overlap between both classes of compounds.

All three models performed well on classifying compounds for both the test and validation sets. However, the RF & MOE, E-state model was chosen as the best model based on the prediction accuracy for the validation set. Further analysis was carried out using this model.

Important descriptors from the Random Forest Model based on MOE & E-state descriptors

We calculated the mean decrease in node impurity as a measure of the descriptor importance. It is the total decrease in node impurities from splitting on the descriptor, averaged over all trees. For classification, the node impurity is measured by the Gini index, which is calculated as follows [13]. Every time a split of a node is made on variable m the Gini impurity criterion for the two descendent nodes is less than the parent node. Adding up the Gini decreases for each individual variable over all trees in the forest gives an easily computed variable importance statistic that is often very consistent with the more complicated permutation importance measure.

The 30 most important descriptors are listed in Table 3 in decreasing order. Hydrophobicity, indicated by SlogP and $\log P_{o/w}$, ranked high. We noticed that hydrophobic compounds tend to be more HLM unstable.

Table 3 Important descriptors from Random Forest model based on MOE and E-state descriptors

Descriptor number	Descriptor	Importance
1	vsa_pol	146.65
2	SlogP	132.56
3	kS_aaCH	118.54
4	logP.o.w.	114.99
5	SMR_VSA2	98.83
6	SMR_VSA5	93.26
7	SlogP_VSA7	88.92
8	kS_dssC	88.57
9	vsa_hyd	87.67
10	TPSA	79.70
11	chi1_C	78.62
12	VDistEq	77.43
13	balabanJ	76.75
14	kC_aaCH	70.63
15	PEOE_VSA.1	70.54
16	PEOE_VSA_HYD	70.51
17	b_ar	69.79
18	PEOE_RPC.	69.44
19	SMR_VSA6	68.25
20	PEOE_VSA.0	67.97
21	kS_ssNH	66.69
22	kS_ssCH2	66.09
23	chi1v_C	65.69
24	PEOE_VSA_FPNEG	64.89
25	kS_aasC	64.65
26	PEOE_PC..1	64.43
27	PEOE_PC.	64.17
28	vsa_acc	63.16
29	KierA2	63.11
30	a_aro	62.77

Interpretation of Bayesian model

An advantage of the ECFP_6 fingerprints is that they are easy to translate into 2D sub-structural sketches within the Pipeline Pilot framework. This allows for a structural analysis of the molecular features that contribute most to HLM stability. Figures 3 and 4 illustrate this. The former shows the ECFP_6 fragments that the naïve Bayesian classifier found most frequently in the stable molecular set. Similarly, Fig. 4 summarizes the most common fragments in the unstable set. By comparing the two figures, it can be seen for example that amide groups in stable compounds are surrounded by non-aromatic cyclic systems, while in unstable compounds those chemical groups are not sterically hindered. If comparisons are extended to the hundreds of fingerprint features obtained for each class of compounds, the atomic environments most favorable for HLM stability can be identified, which constitutes a major advantage in the drug discovery process.

Estimation of prediction confidence

For every prediction model, it is vital to provide a user of the model an estimate of the confidence of the prediction in addition to the prediction itself. We tried several ways to estimate this confidence by calculating the probability of being in one class, the similarity of the query molecule to the closest neighbor in the training set, and the number of nearest neighbors to the query molecule in the training set.

Analysis of prediction accuracy in terms of probability

Using the RF & MOE, E-state model, we analyzed the prediction accuracy in terms of probability in the scale of 0 to 1 for the test set. Table 4 shows the percentage of compounds correctly predicted for each range of probability. As expected, the higher the probability, the higher the percentage of compounds correctly predicted. The analysis shows that, for example, if probability is higher than 0.3, 80% of the compounds are correctly predicted. This information provides a succinct measure of the confidence one can have in the prediction of a query compound.

Prediction confidence in terms of similarity of the nearest neighbor in the training set

We calculated the pair wise similarity of each compound in the test set to each compound in the training set using atom pair similarity described by Carhart. [17]

A neighbor was defined as a compound with similarity higher than or equal to 0.7. Table 5 and Fig. 5 summarize the percentage of compounds correctly predicted for each range of similarity. It is shown that the higher the similarity of the nearest neighbor in the training set, the higher the percentage of compounds correctly predicted. From the user's point of view, prediction confidence can be estimated given the similarity to the nearest neighbor in the training set.

Prediction confidence in terms of number of nearest neighbors in the training set

Table 6 and Fig. 6 compare the percentage of compounds predicted correctly for each range of number of nearest neighbors in the training set. Since we defined a neighbor as a compound with the similarity to the query compound equal to and higher than 0.7 in the training set, there are many compounds in the test set that have no neighbors in the training set. For the rest of the test set, there is at least

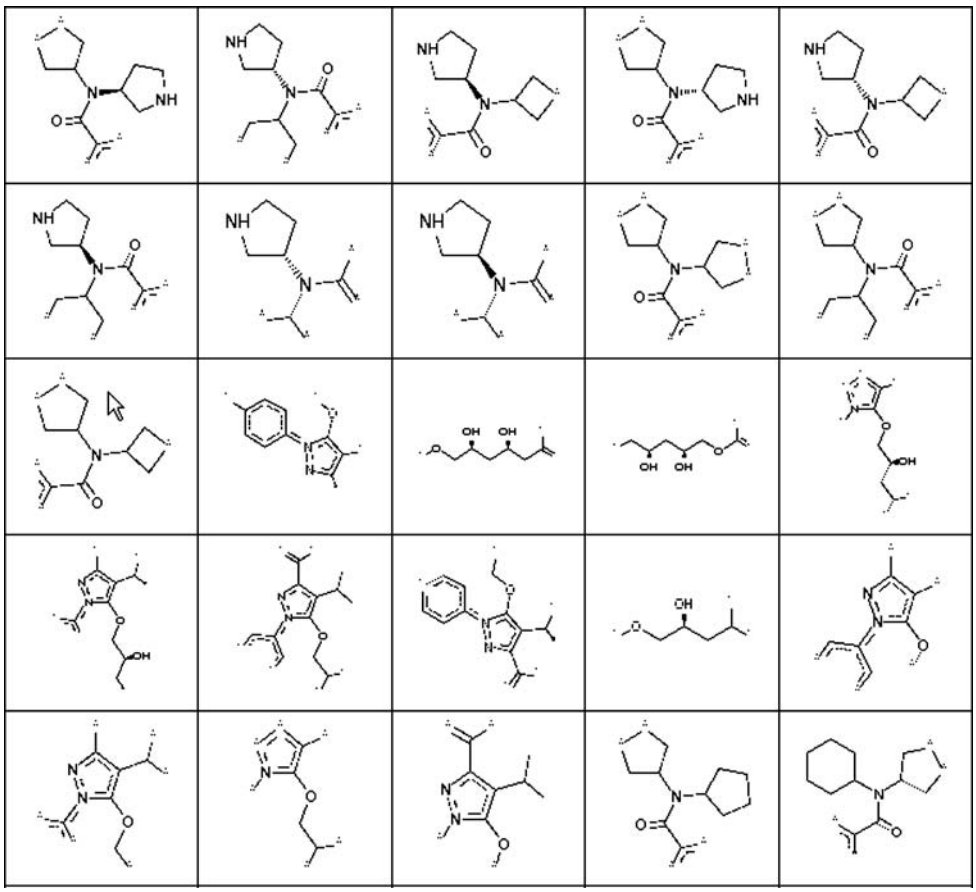

Fig. 3 ECFP_6 sub-structural features most frequently observed in stable compounds. Note: The *stars* in the sketches represent 'any atom'. In some cases, the same ECFP_6 bit is found in several forms (e.g. different isomers or cyclic groups), hence the several 2D representations for the same fingerprint fragment)

one neighbor in the training set. We partitioned the compounds with at least one neighbor into five groups with the similar number of compounds in each group. Generally, the larger the number of neighbors the higher percentage of compounds correctly predicted. The model correctly predicted 72% of the compounds with no neighbors, which is pretty reasonable. However, the percentage of correctly predicted compounds gradually increases as the number of neighbors increases up to 88% for the group with more than 33 neighbors.

From the analysis, the similarity to the nearest neighbor and the number of nearest neighbors in the training set can be used to assess the prediction confidence.

Implementation of the model in a corporate database system

The RF & MOE, E-state model is currently implemented in an internal corporate database system. The model accepts as input: the registration ID, sdf or smiles files, or drawn chemical structures. The model generates the class prediction, (either stable or unstable), a prediction confidence (0–1 scale calculated using class probabilities), the number of nearest neighbors, and the similarity of the closest neighbor.

Application of the model

The model is being used in library design to remove or include compounds for synthesis, to include/exclude compounds for biological testing, and in sub-setting analysis to select representative sets of compounds for actual HLM screening, to reduce the load for in vitro or in vivo screening. Figures 7 to 10 show examples of the application of the HLM model in one of the Pfizer discovery projects. In Fig. 7, the *x*-axis is for the prediction confidence while the *y*-axis is for the experimental value of the apparent intrinsic clearance ($CL_{int,\ app}$) for a set of 1,253 compounds in the project. The compounds predicted to be stable are colored red and the unstable blue. The statistics show that 377 of 465 experimentally stable compounds and the 744 of 788 experimentally unstable compounds are correctly predicted, showing the prediction accuracy of 90%. The Fig. 8 shows the plot of the experimental value versus prediction for a set of 48 compounds within a particular structural series in the project, showing nice separation of the stable from the unstable to be used in virtual library screening. The Fig. 9 shows for a set of 533 compounds within another structural series in the project. In this case, many unstable (false positives) will be retrieved and a few stable

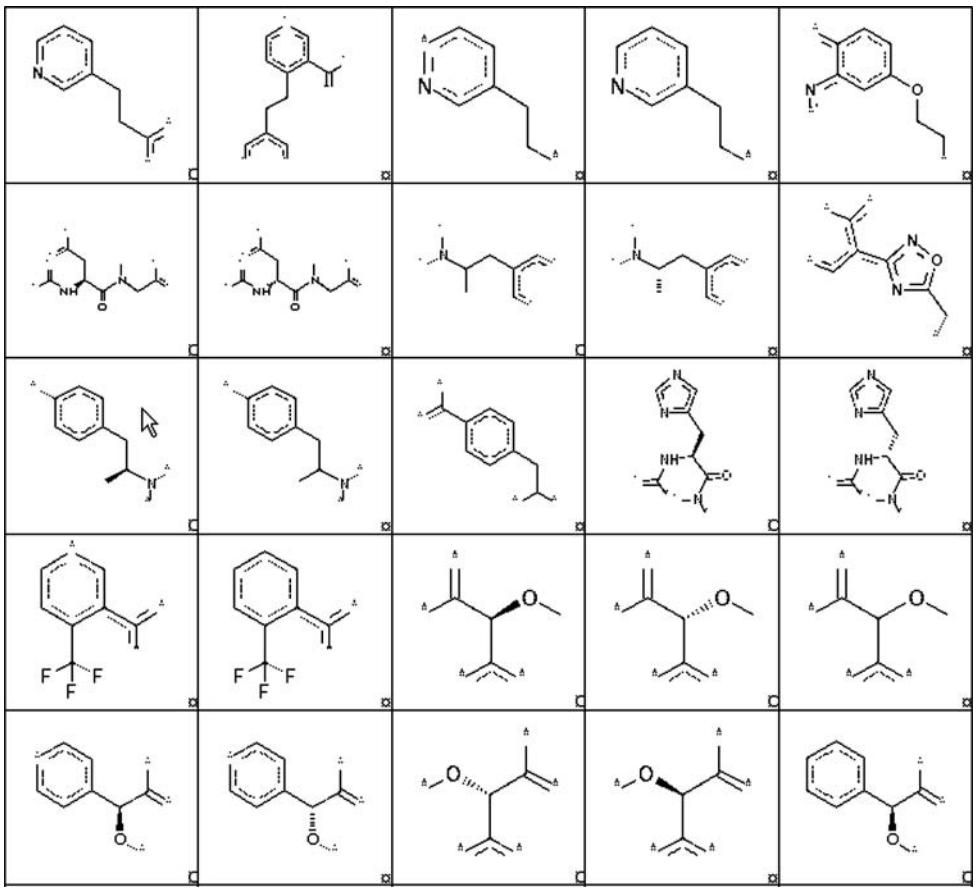

Fig. 4 ECFP_6 sub-structural features most frequently observed in unstable compounds

Table 4 Prediction accuracy in terms of probability of being a stable compound

Probability	Number of cpds	Number of correctly predicted	% Correctly predicted
$0.9 \le P$	257	241	94
$0.8 \le P < 0.9$	240	218	91
$0.7 \le P < 0.8$	224	185	83
$0.6 \le P < 0.7$	240	178	74
$0.5 \le P < 0.6$	239	140	59
$0.4 \le P < 0.5$	277	145	52
$0.3 \le P < 0.4$	300	215	72
$0.2 \le P < 0.3$	354	292	82
$0.1 \le P < 0.2$	412	384	93
$P < 0.1$	368	356	97
Total	2911	2354	81

compounds will be missed. However, model may still be useful for virtual screening if the prediction confidence is higher than 0.8 for stable prediction and lower than 0.2 for unstable prediction. Figure 10 shows the prediction for a set of 343 compounds synthesized after the model was built. The statistics show that 62 of 83 experimentally stable compounds and the 218 of 260 experimentally unstable compounds are correctly predicted, showing the prediction accuracy of 82%. In this case, if considered only for compounds with prediction confidence higher than 0.6 for stable ($N = 42$) and lower than 0.4 for unstable ($N = 169$), the prediction accuracy increases to 94%.

We recommend the users to assess the applicability of the model in their projects using the prediction confidence before applying it.

Table 5 Prediction confidence in terms of similarity of the nearest neighbor in the training set

Similarity	Number of cpds	Number of corr	% Corr	Number of incorr	% Incorr
0	239	173	72	66	28
$0.7 < \text{Sim} \le 0.8$	434	302	70	132	30
$0.8 < \text{Sim} \le 0.9$	1189	975	82	214	18
$0.9 < \text{Sim} \le 1.0$	1049	904	86	145	14

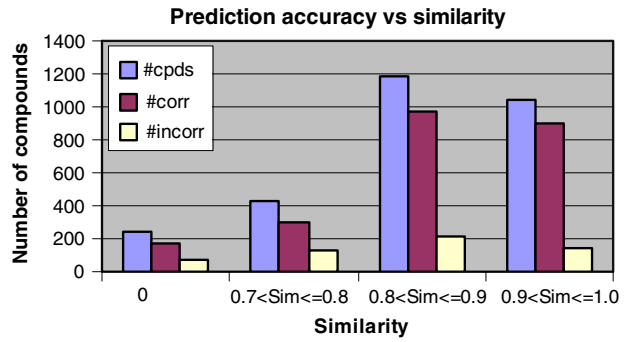

Fig. 5 Prediction accuracy in terms of similarity to the nearest neighbor in the training set

Table 6 Prediction confidence in terms of number of nearest neighbors in the training set

Number of neighbors	Number of cpds	Number of corr	% Corr	Number of incorr	% Incorr
0	239	173	72	66	28
$1 \leq NN \leq 3$	571	433	76	138	24
$4 \leq NN \leq 8$	556	431	78	125	22
$9 \leq NN \leq 16$	539	455	84	84	16
$17 \leq NN \leq 32$	533	445	83	88	17
$NN > 33$	473	417	88	56	12

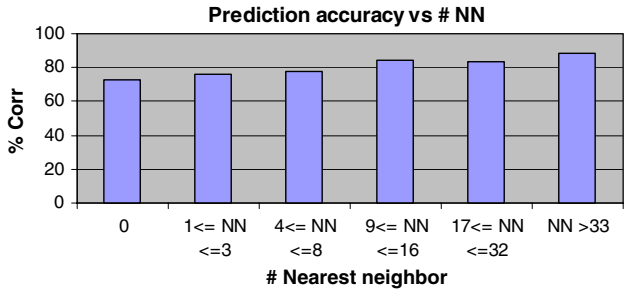

Fig. 6 Prediction confidence in terms of number of the neighbors in the training set

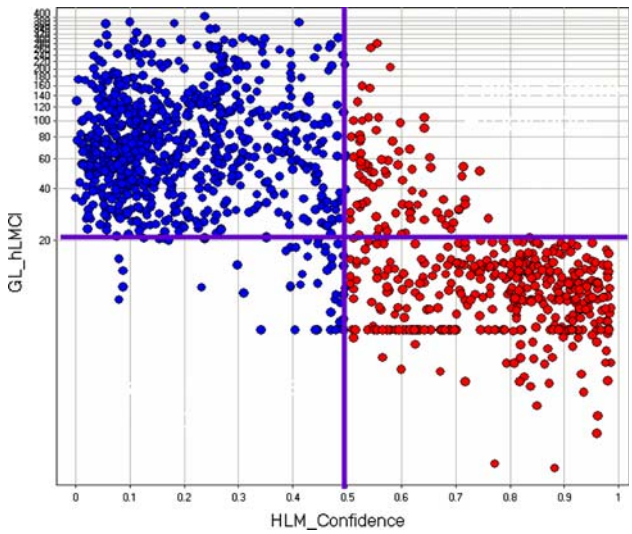

Fig. 7 HLM prediction for a Insert discovery project

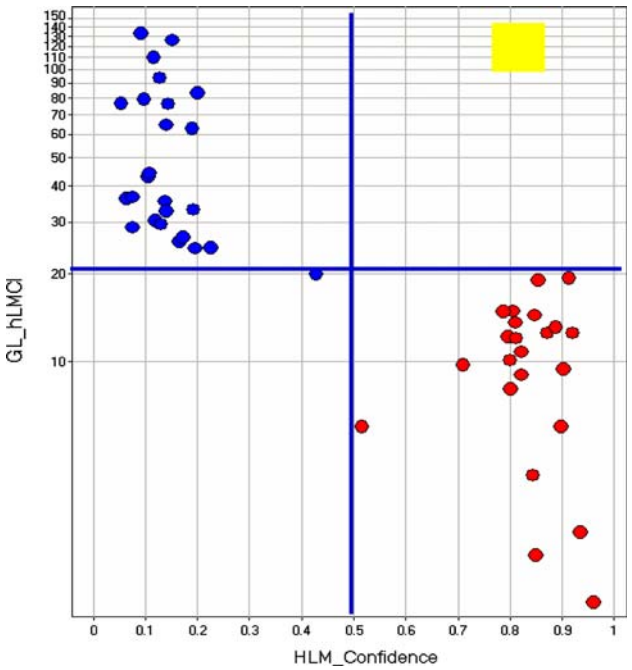

Fig. 8 HLM prediction for one structural series within a Pfizer discovery project

Conclusion

We developed highly predictive classification models for (HLM) stability using the apparent *intrinsic* clearance ($Cl_{int,\ app}$) as the end point. Using Random Forest and Bayesian classification methods with MOE, E-state descriptors, ADME Keys, and ECFP_6 fingerprints, various highly predictive models were developed. The best performance of the models shows 80 and 75% prediction accuracy for the test and validation sets, respectively. From the detailed analysis of prediction accuracy for the test set in terms of probability, similarity to the nearest neighbor, and the number of neighbors in the training set, the model provides multiple estimates of the confidence in the prediction. Analysis of significant descriptors shows the molecular hydrophobicity is an important descriptor to differentiate the HLM stable from unstable compounds.

The naïve Bayesian classifier based on the ECFP_6 fragments identified structural fragments frequently found in the groups of HLM stable compounds and unstable compounds, which can be used in the design of new drugs.

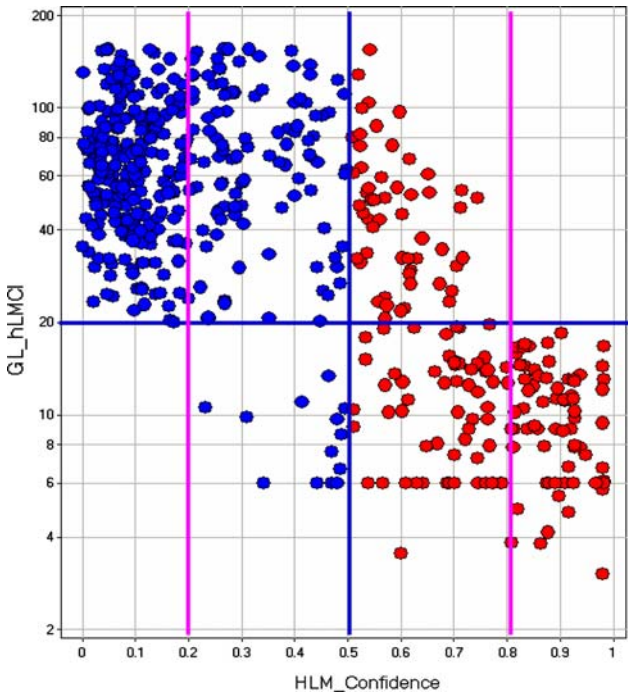

Fig. 9 HLM prediction for one structural series within a Pfizer discovery project

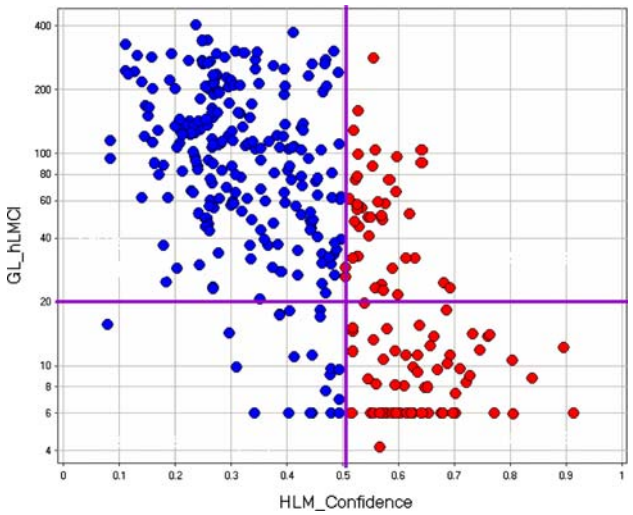

Fig. 10 HLM prediction for a Pfizer discovery project after model was built

The model is currently being used in many useful ways in drug discovery projects.

Acknowledgment We would like to thank Rob Goulet and Shao-Tien Sng for their technical support to implement the model in Rgate, Genevieve Paderes and Klaus Dress for providing use cases, Cornel Catana, Ben Burke, Meihua Tu, Jason Hughes, Chad Stoner, Marcel de Groot, and Eric Gifford for their scientific discussions. Special thanks goes to Dan Ortwine for the critical reading of the manuscript.

References

1. Kwon Y (2001) Handbook of essential pharmocokinetics, pharmacodynamics and drug metabolism for industrial scientists, 1st edn. Springer, Berlin
2. Jolivette LJ, Ekins S (2007) Adv Clin Chem 43:131
3. Taylor C, Janisewski J, Whalen K (2006) Integrated high-throughput ADME: a 384-well human liver microsome assay to determine metabolic profiles of 1800 compounds per week. Poster, AAPS. Annual Meeting
4. JMP 5.1.1, SAS Publishing, SAS Campus Drive, Cary, NC 27513–2414
5. Molecular Operating Environment (MOE) software 2005.06, Chemical Computing Group Inc., 1255 University Street, Montreal, Quebec, Canada, H3B 3 × 3
6. Hall LH, Kier LB (1999) Molecular structure description: the electrotopological state. Academic, New York
7. The ADME keys were coded by Jing Lu from different resources: The majority of the ADME keys were not from literature. Only a small portion was from *J. of Chemical and Computer Information Science* 1997, 37, 329. The other sources include common functional groups and rings liable to metabolism based on personal understanding, ISIS Keys, Hetero rings from John Blake, A few fragments from Pfizer at Nagoya, Japan
8. ISIS/Base, Version 2.1.3; Molecular Design Ltd. (14600 Catalina Street, Irvine, CA 92714
9. Pipeline Pilot 4.5 is a program of SciTegic (Accelrys, Inc.), 10188 Telesis Court. Suite 100, San Diego, CA 92121–4779 USA (http://www/scitegic.com/)
10. Rogers D, Brown RD, Hahn M (2005) J Biomol Screen 10(7):682
11. Morgan HL (1965) J Chem Doc 5:107
12. Brieman L (2001) Mach Learn 45:5
13. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP (2003) J Chem Inf Comput Sci 43:1947
14. R Development Core Team (2004) R: a language and environment for statistical computing. http://www.R-project.org
15. Migliavacca E (2003) Mini-Rev Med Chem 3:831
16. Dunn G, Everitt B (1995) Clinical biostatistics: an introduction to evidence-based medicine. Edward Arnold, London
17. Carhart RE, Smith DE, Venkataraghavan R (1985) J Chem Inf Comput Sci 25:64