Towards Computational Prediction of Biopharmaceutics Classification System: A QSPR Approach *

Hai Pham-The [1,*], Huong Le-Thi-Thu [2], Teresa Garrigues [3], Marival Bermejo [4], Isabel González-Álvarez [4] and Miguel Ángel Cabrera-Pérez [3,4,5]

[1] Hanoi University of Pharmacy, 13-15 Le Thanh Tong, Hoan Kiem, Hanoi, Vietnam
[2] School of Medicine and Pharmacy, Vietnam National University, 144-Xuan Thuy, Cau Giay, Hanoi, Vietnam; E-Mail: ltthuong1017@gmail.com
[3] Department of Pharmacy and Pharmaceutical Technology, University of Valencia, Burjassot 46100, Valencia, Spain; E-Mails: Teresa.Garrigues@uv.es (T.G.); macabreraster@gmail.com (M.A.C.-P.)
[4] Department of Engineering, Area of Pharmacy and Pharmaceutical Technology, Miguel Hernández University, 03550 Sant Joan d'Alacant, Alicante, Spain; E-Mails: mbermejo@umh.es(M.B.)
[5] Unit of Modeling and Experimental Biopharmaceutics, Chemical Bioactive Center, Central University of Las Villas, Santa Clara, 54830, Villa Clara, Cuba
* Author to whom correspondence should be addressed; E-Mail: thehai84@yahoo.com; Tel.: +84-996-888-868; Fax: +84-4-39710550.

Received: 29 October 2015 / Accepted: 29 October 2015 / Published: 2 December 2015

Abstract: Today classification of drug candidates on the Biopharmaceutics Classification System (BCS) has become an important issue in pharmaceutical researches. In this work, we provide a potential *in silico* approach to predict this system using two separately classification models of Dose number and Caco-2 cell permeability. 18 statistical linear and nonlinear models have been constructed based on 803 0-2D Dragon and 126 Volsurf+ molecular descriptors to classify the solubility and permeability properties. The voting consensus model of solubility (VoteS) showed a high accuracy of 88.7% in training and 92.3% in test set. Likewise, for the permeability model (VoteP), accuracy was 85.3% in training and 96.9% in test set. A combination of VoteS and VoteP appropriately predicts the BCS class of drugs (overall 73% with class I precision of 77.2%). This consensus system predicts the BCS allocations of 57 drugs appeared in the WHO Model List of Essential Medicines with 87.5% of accuracy. A simulation of a biopharmaceutical screening assay has been proved in a large data set of 37,377 compounds in different drug development phases (1, 2, 3 and launched), and NMEs. Distributions of BCS forecasts illustrate the current status in drug discovery and development. It is anticipated that developed QSPR models could offer the best estimation of BCS for NMEs in early stages of drug discovery.

Keywords: Biopharmaceutics Classification System (BCS); Dose Number; Caco-2 cell permeability; Quantitative Structure Activity/Property Relationship (QSAR/QSPR)

Mol2Net YouTube channel*: http://bit.do/mol2net-tube*
*The full content of this communication can be partly found in Pham-The H *et al*. Mol. Pharmaceutics, 2013; 10(6): 2445-61 and Pham-The H *et al*. Mol. Inf. 2013; 32(5-6): 459-79.

1. Introduction

After almost 20 years of the introduction and exploration of the Biopharmaceutics Classification System (BCS), it has gained a major impact on the regulation and development of immediate release (IR) solid oral drug products [1,2]. Based on the principal factors that determine the rate and extent of drug absorption, the BCS provides a scientific framework for classifying drug substances into one of four categories. According to BCS, IR solid oral dosage forms are categorized as having either rapid or slow *in vitro* dissolution, and then classified based on aqueous solubility and intestinal permeability of the active pharmaceutical ingredient (API) [1]. This system has been formally adopted by the US FDA [3], the European agency EMEA [4] and the World Health Organization (WHO) [5] as a technical standard for waiving BE test requirements for oral drugs. A recent study of the economic impact of granting biowaivers for class I and III BCS demonstrated an impressive saving annual expenditure on running BE studies, being more than 120 million dollars between the two classes [6]. Because it avoids unnecessary drug exposures to healthy subjects, while maintaining the high public health standard for therapeutic equivalence, the BCS is, without doubt, a potential tool for speeding up and reducing the cost of drug development.

There is a continuing effort worldwide to detect, in the early discovery, the possible BCS-based biowaiver candidates, e.g. BCS class I drugs [7]. One of the common strategies is based on BCS provisional classification in which the drugs are classified by two sources: dose related solubility data (Dose number, Do) and estimated human absorption data, i.e. *in vitro* permeability (usually determined by the Caco-2 cell cultured method) [3,8], or simple *in silico* partition coefficient calculation [9]. In this regard, *in silico* approach presents the two most important advantages: (i) provides a flexible approach that can be applied in different stages of drug development with different purposes, and (ii) allows estimating the BCS classes of new molecular entities (NMEs) without knowledge of therapeutic dosage. Definitely, with respect to experimental methods, computational approaches are cost-saving and no sample requirement methods.

However, up to now, robust *in silico* approach, i.e. Quantitative Structure-Activity/Property Relationships (QSAR/QSPR) modeling, has not been explored sufficiently in the BCS studies. Based on published findings [10], and to respond to the rising need of early identification of possible biowaiver drugs, in this work, we attempt to develop robust QSPR models to classify the solubility and permeability terms that compose the BCS (Figure 1). These models were rigorously validated on various published BCS class drug sets [5,9,11-13] and the feasibility of performing PBC prediction in early drug discovery is discussed.

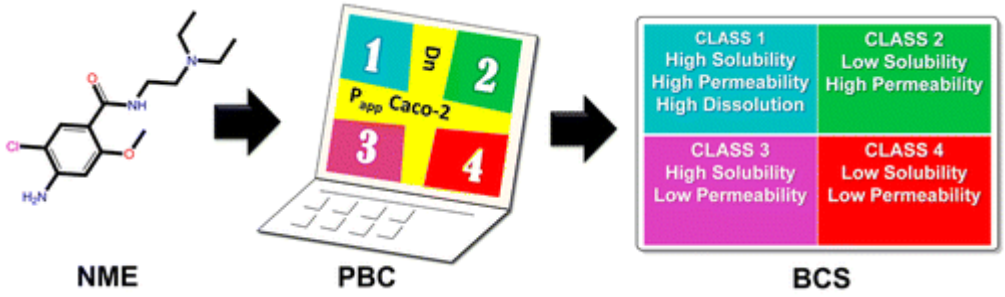

Figure 1. Summary scheme of current in silico study

2. Results and Discussion

In 2004, a number of 123 orally administered drugs on the World Health Organization (WHO) Essential Medicine List (EML) were initially classified into BCS [9,11]. Later, 200 oral drug products in the United States, Great Britain, Spain, and Japan were classified based on published solubility data and permeability data estimated by calculated log *P* [12]. Recently, increasing attention has been turned out for determining the Provisional Biopharmaceutical location of orally administered immediate-release (IR) drug products using different estimated gastrointestinal permeability, such as partition coefficients (log *D* and log *P*), molecular surface area (PSA) or other *in vitro* permeability.[14-17] It has been emphasized that the distribution of BCS class I, II, III, and IV in each classification are quite different. In this report, taking advantage of the availability of experimental *in vitro* Caco-2 cell data a Provisional Biopharmaceutical Classification (PBC) of 322 oral drug products gathered from literature was performed. To our knowledge, it is the largest data set for such classification. Classifications of current data are described in [7].

Physicochemical profiling of PBC. It is very useful to analyze the *similarity* between physicochemical spaces characterized by PBC classes, especially for developing computational predictions of current PBC and further BCS. Thus, six commonly used physicochemical parameters were calculated by Dragon and Volsurf+ for this analysis:[18,19] molecular weight (MW), polar surface area (PSA), Mlog *P*, log D_6, log $D_{7.5}$, total number of hydrogen bond donors and acceptors (nHA+B), number of free rotatable bonds (RBN), and estimated ionization states. The average and median values of maximum dose strength (D_{max}) as well as Caco-2 P_{app} were also analyzed for each class.

Unsurprisingly, class II drugs display the highest lipophilicity, while class III and IV are more hydrophilic. Class I drugs represent a balanced physicochemical profile even though they tend to be more lipophilic. In general, only the hydrogen bonding term is fairly different from one class to another. There is certain *physicochemical similarity* between class I and II (Mlog *P*, log *D* at basic medium), class III and IV (nHA+B, PSA), or class II and III (MW), etc. Values of D_{max} do not present any trend. It is demonstrated that poor bioavailability is more likely when the compounds violate two or more of the Lipinski's rules (Ro5): (i) log *P* <5, (ii) MW< 500, (iii) HBD (hydrogen bond donors) < 5, and (iv) HBA (hydrogen bond acceptors) < 10.[20] Current data was collected mostly among successful drugs. Then, it is easy to understand that many of them (>95%) passed the Ro5.

Computational models to predict PBC class from chemical structures. Solubility and Caco-2 permeability were modeled independently. The final computational PBC classification was

achieved using two voting consensus (permeability and solubility) systems. QSPR models obtained by different statistical techniques for each property are described below.

Solubility modeling. Three model series were obtained using LDA, QDA and BLR. Different molecular descriptors (MDs) were used for building QSPR models. From every model series constructed with every technique, the best one was selected (detailed comparisons are described in supplement documents). Table 1 summarizes the mathematical equations and performances of the three best models for PBC solubility prediction.

Permeability modeling. The same procedure was carried out to select the best classifiers for PBC permeability class. Table 2 displays the relevant information of permeability models.

Classifications of four PBC classes. The two obtained voting models were finally combined to estimate the four PBC classes of the data (322 compounds). Table 3 displays the confusion matrix of this consensus system. A good overall accuracy of 73.0 % was obtained by this system..

Analysis of molecular descriptors (MDs). Interestingly, the PBC solubility and permeability terms are well described using a small set of MDs.

Table 1. Performances of the three best models for PBC solubility classification

Technique	Descriptor family	MCC	Accuracy	Specificity	Sensitivity	Precision	AUC (Ts)[b]
		% (Tr/Ts)[a]					
LDA (S1)	0-2D Dragon *plus* Volsurf+	0.66/0.54	83.3/76.9	82.2/79.3	84.0/75.0	86.9/81.8	0.88±0.04
QDA (S2)	0-2D Dragon	0.63/0.75	81.7/87.7	82.2/82.8	81.3/91.7	86.5/86.8	0.97±0.04
BLR (S3)	0-2D Dragon	0.60/0.69	80.5/84.6	75.5/82.1	84.1/86.5	83.0/86.5	0.96±0.03
VoteS	All	**0.68/0.87**	**84.4/93.9**	**85.0/89.3**	**84.0/97.2**	**88.7/92.3**	–

Mathematical equations

$CLASS_{Do}(+/-) = -1.59 - 0.54\times P\_VSA\_v\_3 + 0.80\times nArC{=}N + 0.65\times C\text{-}005 - 0.84\times CATS2D\_04\_AL + 0.79\times DLS\_04 + 4.51\times ID3 + 0.28\times A - 0.41\times LgD5$ **(S1)**

$N = 257 \quad \lambda = 0.60 \quad D^2 = 2.74 \quad F = 25.61 \quad p < 0.0001$

$CLASS_{Do}(+/-) = -0.36 - 0.90\times Me - 1.40\times nCt - 0.79\times NssNH + 1.22\times BLTD48 + 0.87\times DLS\_04 - 0.82\times CMC\text{-}50 - 1.86\times nArC{=}N\times N\text{-}067 + 0.41\times N\text{-}067\times NssNH - 0.73\times Me\times CMC\text{-}50 + 0.51\times nR10^2$ **(S2)**

$N = 257 \quad \lambda = 0.59 \quad D^2 = 2.88 \quad p < 0.0001$

$Ln(P+/P-) = 2.63 - 0.59\times nCp + 4.44\times nArC{=}N + 0.20\times H\text{-}052 + 1.82\times N\text{-}067 - 1.32\times NssNH + 1.09\times BLTD48 + 4.58\times LDS\_04 - 1.38\times CMC\text{-}50 - 0.38\times nO$ **(S3)**

[a]Measured performances of training/test set; [b]Area under the ROC curve determined on test set by non-parametric assumptions in 95% asymptotic confidence interval.

Table 2. Performances of the three best models for PBC permeability classification

Technique	Descriptor family	MCC	Accuracy	Specificity	Sensitivity	Precision	AUC (Ts)[b]
		% (Tr/Ts)[a]					
LDA (P1)	0-2D Dragon *plus* Volsurf+	0.63/0.69	81.6/84.9	81.9/85.7	81.4/84.2	82.0/88.9	0.93±0.03
QDA (P2)	0-2D Dragon	0.65/0.76	82.4/87.9	81.1/89.3	83.7/86.8	81.8/91.7	0.94±0.03
BLR (P3)	0-2D Dragon *plus* Volsurf+	0.64/0.73	82.0/86.4	79.5/89.3	84.5/84.2	80.7/91.4	0.92±0.03
VoteP	All	**0.70/0.77**	**85.2/87.9**	**85.0/96.4**	**85.3/81.6**	**85.3/96.9**	–

Mathematical equations

$CLASS_{Papp}(+/-) = -5.91 + 0.01\times P\_VSA\_s\_6 - 1.62\times nRNR2 - 0.74\times C\text{-}016 + 2.64\times CATS2D\_08\_AP + 4.23\times LLS\_01 + 0.01\times WN2 + 3.79\times CACO2$ **(P1)**

$N = 256 \quad \lambda = 0.57 \quad D^2 = 2.81 \quad F = 22.24 \quad p < 0.0001$

$CLASS_{Papp}(+/-) = 0.32 - 1.02\times GATS2m + 0.95\times GATS2s - 0.55\times nRNR2 - 0.52\times B03[O\text{-}O]$ **(P2)**

$$- 1.95\times SAdon + 0.82\times LLS\text{-}01 + 3.46\times nC{=}N\text{-}N{<}\times B04[O\text{-}Cl] + 0.37\times nRNR2\times SAdon + 0.32\times CATS2D\_03\_DD\times SAdon - 0.46\times B08[C\text{-}O]^2$$

$N = 256 \quad \lambda = 0.55 \quad D^2 = 3.18 \quad p < 0.0001$

$$Ln\,(P+/P-) = 5.49 - 2.05\times nRNR2 + 3.74\times CATS2D\_07\_DP + 1.88\times CACO2 - 5.04\times GATS2m - 22.48\times nFuranes - 0.02\times SAdon - 1.05\times nRCOOH \quad \textbf{(P3)}$$

[a]Measured performances of training/test set; [b]Area under the ROC curve determined on test set by non-parametric assumptions in 95% asymptotic confidence interval.

It is important to note that there are some MDs directly related to polarizability and dispersion forces within molecules (*nCp*, *nCt*), molecular size (*nR10*, P_*VSA*_*v*_*3*), lipophilicity and hydrophobicity (*BLTD48*, *CATS2D_04_AL*, *CMC-50*), and especially, the polar, chargeable and hydrogen bond forming capacity (*A*, *Me*, *nO*, *nArC=N*, *C-005*, *N-067*, *NssNH*, *LgD5*). Beside, rule based MDs, which represent common physicochemical combination trends of known drug-like and lead-like dataset,[21,22] are selected. Generally, current finding structure-property (*Do*) relationship (S*Do*R) are rather similar with Khandelwal *et al*.'s analysis.[23]

On the other hand, the ionization state (*GATS2s*, *P_VSA_s_6*, *nRNR2*), molecular size (*GATS2m, nFuranes*, *C-016*) and hydrogen bond donor and acceptor regions (*nRCOOH*, *nRNR2*, *nC=N-N<*, *CATS2D_03_DD*, *CATS2D_07_D*, *CATS2D_08_AP*, *SAdon*, *WN2* etc.) are well correlated with Caco-2 permeability. The ADME descriptor *CACO2* was selected two times in permeability models. Please note that numeric values of this variable are result of partial least square (PLS) discriminant analysis developed by Zamora *et al*.[24] Unfortunately, the use of this descriptor does not provide precise knowledge of descriptor impacts on PBC permeability class.

Regulatory validation and applications of in silico PBC models. A robust forecast of PBC class is very useful in early drug discovery. Especially, for many NMEs whose therapeutic dose-ranges are not available in preclinical stages. This is also important for estimating possible BCS memberships, since there is a great correspondence between proposed PBC and BCS cited in regulatory guidelines [5].

Table 3. Confusion matrix of consensus system for the prediction of PBC classes

	Predicted PBC Class I	Predicted PBC Class II	Predicted PBC Class III	Predicted PBC Class IV	Total	Accuracy (%)	MCC
PBC Class I	**61**	11	18	1	91	67.0	0.62
PBC Class II	10	**59**	2	5	76	77.6	0.67
PBC Class III	7	4	**74**	12	97	76.3	0.63
PBC Class IV	1	8	8	**41**	58	70.7	0.63
Total	79	82	102	59	322		
Precision (%)	77.2	72.0	72.5	69.4			

Biopharmaceutical Screening Simulations

Finally, a large database of drugs, clinical and non-clinical trial compounds was subjected to computational prediction using *in silico* PBC consensus model. A total number of 37,202 compounds were analyzed (Figure 2). Recently, this database was classified by *in silico* BDDCS consensus models to estimate the distribution of BDDCS class.[10] In contrast to that study, obtained models here are employed for comparing the predictions and then making a round estimation of the distribution of BCS class. It is important to note that some compounds obtained non-conclusive-classification due to

their condition of outliers of Ads. 1699 compounds (4.6% of prediction data) are classified as I/II, I/III, II/IV, and III/IV. Most of them (1512 compounds) are low-activity (W6) and high-activity (W9) compounds.[10] Especially, 29 compounds could not be classified by *in silico* models. Among those conclusively predicted as PBC class I, II, III and IV, there exists similar proportion between launched and clinical phase 3 drugs, between clinical phases 1 or 2 drugs and W6 or W9 compounds.

As can be appreciated from Figure 3, more than 40% of drugs and phase 3 are similar to PBC class I. The phase 3 compounds similar to PBC class II significantly overcome the PBC class III but for drugs, their percentage become similar. Compounds classified as PBC class IV take the minimal proportion in the two drug sets (7-8%). In contrast, about 50% of phase-1 and phase-2 drugs are predicted as PBC class II. This percentage is even greater (62-63%) in W6 and W9 datasets. Compounds predicted as PBC class I maintain the same proportion with respect to phase 1, 2, W6 and W9 whole dataset. There is a noticeable change of the predicted PBC class III for phase 1 and 2 drugs (15-18%) compared to W6 and W9 (7%) compounds. Particularly, compounds of W6 data set, predicted PBC class IV compounds outnumber those of predicted as PBC class III. These trends of PBC class predictions reflect the drug development process and agree, in turn, upon some points with previous findings.[10,25]

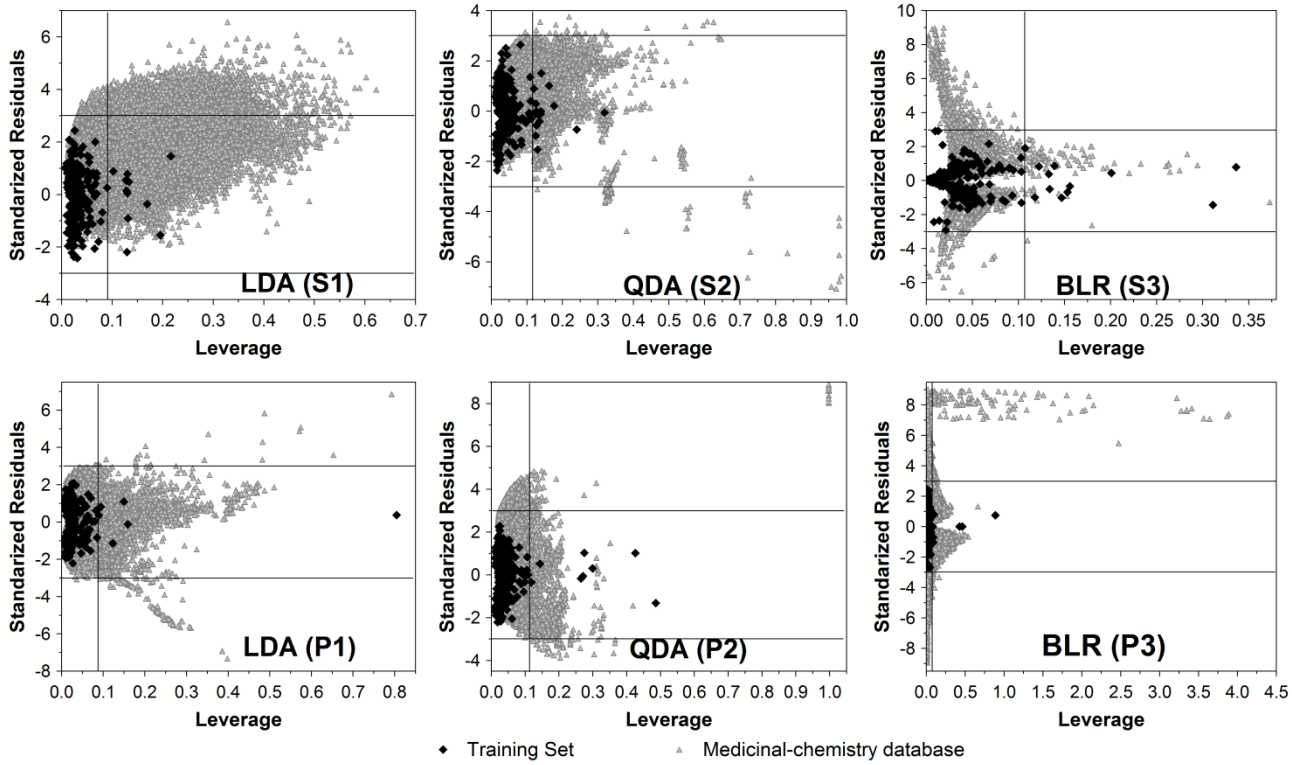

Figure 2. William's plots based on solubility and permeability models for training and screening large medicinal-chemistry database

3. Materials and Methods

Data set. BCS based-provisional classification requires both solubility and permeability measurements. In this work, a set of 322 drugs was obtained from published works. A provisional classification was executed by means of an extensive literature revision of experimental values and assigned classes, as follows.

Solubility data. The drug solubility data (in mg/mL) can be obtained from standard references,[9] such as the Pharmacopeia [26] or the Merck Index.[27] Due to the extensive survey, herein we only report the lowest solubility under the conditions listed above. In addition, scale-up guidelines were taken from Kasim *et al*. whenever solubility data was not available or was undefined.[9]

Maximum Dose Strength. Two reference sources were mainly used for searching values of maximum dose strength (mg): (i) the WHO Model List of Essential Medicines,[28] and (ii) Orange Book.[29] For drugs that are not included in these documents or exist in different market presentations, the first introduced strengths were revised and used as highest dosages. Doses in mg/kg were transformed into mg assuming 70Kg as body weight.

*Dose Number Calculations***.** The dose number (D_0) was calculated using the following equation

$$D_0 = \frac{(M0/V0)}{S} \quad (1)$$

where, M_0 is the highest dose strength (mg), S is the aqueous solubility (mg/mL) under conditions mentioned above and water volume V_0 is assumed to be 250 mL.[1,9] Drugs with $D_0 \leq 1$ were classified as high-solubility drugs. Conversely, drugs with $D_0 > 1$ were assigned as low solubility drugs.[9]

Permeability Estimations. In this work, *in vitro* Caco-2 cell permeability is used to classify drug according to BCS. For this purpose, we take advantage of our previous research where an extensive literature survey of this kind of data was processed.[30] Besides, we have adopted the same method proposed by Kim *et al*.,[31] taking the average permeability value of Metoprolol (average apparent permeability P_{app} = 20×10^{-6} cm/s) for benchmarking the high permeability class boundary. Due to the large revised literature, the mean values were listed, excluding those laid outside of the mean±2SD (standard deviation) ranges. Additionally, available data obtained on both directions apical to basolateral ($P_{app,\ A\text{-}B}$) and viceversa ($P_{app,\ B\text{-}A}$) were taken into account.

Computational methods. Taking all above together, in this work efforts have been made to establish really useful statistical predictors for BCS classes of NMEs based on two separate model series of dose number and Caco-2 cell permeability. To attain this purpose, the following computational procedures should be considered: (i) suitably computing physicochemical and molecular descriptors, (ii) rational selection of training and test sets, (iii) establishment of modeling strategy and appropriated variable selection, and (iv) ascertainment of BCS predictions for NMEs in the context of regulatory statements.

Molecular descriptor calculations. 803 simple (0-2D) descriptors belonging to 29 families implemented in Dragon software *version* 6.0,[19] and 126 molecular descriptors in VolSurf+ *version* 1.0.4 [18] were calculated.

Model building and feature selection. Three statistical classification algorithms were applied in order to detect all possible (linear or non-linear) relationships between solubility/permeability and computed parameters: LDA (Linear Discriminant Analysis), QDA (Quadratic Discriminant Analysis) and BLR (Binary Logistic Regression).

Performances of models were evaluated using false positive rate (FPr), true negative rate (TN, for specificity), true positive rate (TP, for sensitivity), Matthews Correlation Coefficient (MCC) and predictive accuracy, as defined below:

$$\text{Specificity} = TN/(TN+FP) \quad (2)$$

$$\text{Sensitivity} = TP/(TP+FN) \quad (3)$$

$$\text{Precision} = TP/(TP+FP) \quad (4)$$

$$MCC = [(TP\times TN) \times (FP\times FN)]/[(TP+FP)(TP+FN)(TN+FP)-(TN+FN)]^{1/2} \quad (5)$$

$$\text{Accuracy} = (TN+TP)/(TN+TP+FN+FP) \quad (6)$$

For reliable predictions of these three *external* datasets, it is important to consider all applicability domains (ADs) defined by the chemical spaces of the training set. There are many approaches for AD estimation.[32] Here, the leverage approach, a geometric method commonly used for QSAR problems, was employed. The leverage of a compound in the original variable space is defined as $h_i = [X(X'X)^{-1}X']$, where X is the descriptor matrix derived from the training set descriptor values. The warning leverage (h*) is defined as $h^*=3(p+1)/n$, where n is the number of training compounds, and *p* is the number of predictor variables [32]. Compounds with $h_i > h^*$ were observed to reveal their influence on classification performance. It is not necessary to exclude them from predictions although they appear to be outside AD. However, compounds are considered to be outliers if they lay outside the ±3 standardized residual (δ) range [32].

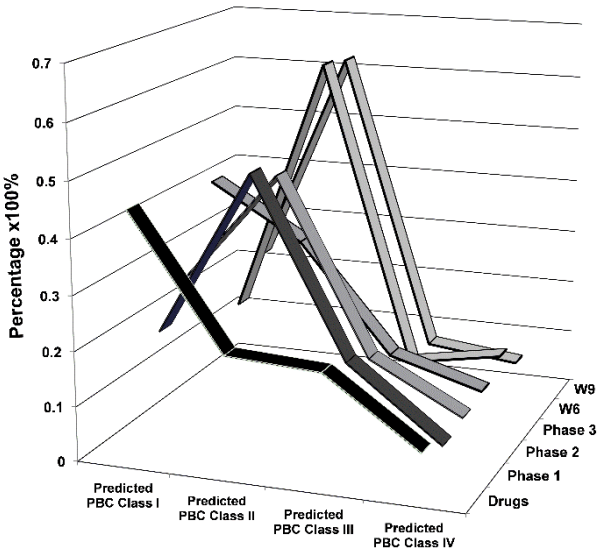

Figure 3. . Distribution comparison of computational PBC assignments of launched drugs, compounds in different drug development stages (phase 1, 2, 3), and bioactive micromolar (W6) and nanomolar (W9) compounds.[10]

4. Conclusions

In this report, a systematic study was carried out in order to standardize a BCS-based provisional classification of 322 drugs and develop computational predictions of BCS class for NMEs. It is of great interest to assign as soon as possible the probable BCS class of a drug candidate. By using extensively revised references of solubility and *in vitro* Caco-2 permeability, a very commonly used preclinical assay in pharmaceutical industry, a better *in vivo* BCS classification of drugs is anticipated. Consequently, the classification results in this study display a high concordance with BCS classification of common regulatory authorities (WHO, FDA). Other classification schemes were compared with PBC. Large additional information concerning the BCS classification of current data was analyzed in order to identify advantages as well as limitations when using PBC. As an attempt to develop QSPR models able to predict the PBC class, it was demonstrated the possibility of screening NMEs in the early phase of drug development.A combination of *in silico* and *in vitro* approaches provides a basis for robust estimation of the BCS class of NMEs without clinical information and contribute to early selection of biopharmaceutical promissory drug candidates. As a relevant limitation, this data set consists of a small number of drugs. Besides, the uncertainty of the relationship between absorption extent and proposed provisional classification (especially for low absorbed drugs) remains. A modification of BCS classification scheme (particularly for

class II and III) is needed. A further compilation of *in vitro* permeability data and aqueous solubility may enhance the applicability domain of *in silico* classifications.Main text paragraph.

Acknowledgments

H.L-T-T is supported by Vietnam National University. H.P-T, M.B, I.G-A, T.G and M.A.C-P acknowledge financial support of AECID (Grant No. 1- D/031152/10 and DCI-ALA/19.09.01/10/21526/245-297/ALFA 111(2010)29.

Author Contributions

All the authors contributed equally.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. Amidon, G.L.; Lennernas, H.; Shah, V.P.; Crison, J.R. A theoretical basis for a biopharmaceutic drug classification: The correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm. Res.* **1995**, *12*, 413-420.
2. Chen, M.L.; Amidon, G.L.; Benet, L.Z.; Lennernas, H.; Yu, L.X. The bcs, bddcs, and regulatory guidances. *Pharm. Res.* **2011**, *28*, 1774-1778.
3. CDER/FDA. *Fda guidance for industry: Waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system*; Federal Drug and Food Administration: Rockville, MD, USA: Center for Drug Evaluation and Research, 2000.
4. CPMP/EWP/QWP/1401/98. *Note for guidance on the investigation of bioavailability and bioequivalence*; The European Agency for the Evaluation of Medicinal Products (EMEA): London, December 14, 2000.
5. *Annex 8: Proposal to waive in vivo bioequivalence requirements for who model list of essential medicines immediate-release, solid oral dosage forms*; Technical Report Series No. 937; WHO Expert Committee on Specification for Pharmaceutical Preparations: 2006; pp 391-461.
6. Cook, J.A.; Davit, B.M.; Polli, J.E. Impact of biopharmaceutics classification system-based biowaivers. *Mol. Pharmaceutics* **2010**, *7*, 1539-1544.
7. Pham-The, H.; Garrigues, T.; Bermejo, M.; González-Álvarez, I.; Monteagudo, M.C.; Cabrera-Pérez, M.Á. Provisional classification and in silico study of biopharmaceutical system based on caco-2 cell permeability and dose number. *Mol. Pharmaceutics* **2013**, *10*, 2445-2461.
8. Dahan, A.; Lennernäs, H.; Amidon, G.L. The fraction dose absorbed, in humans, and high jejunal human permeability relationship. *Mol. Pharmaceutics* **2012**, *9*, 1847−1851.
9. Kasim, N.A.; Whitehouse, M.; Ramachandran, C.; Bermejo Sanz, M.; Lennernas, H.; Hussain, A.S.; Junginger, H.E.; Stavchansky, S.A.; Midha, K.K.; Shah, V.P., *et al.* Molecular properties of who essential drugs and provisional biopharmaceutical classification. *Mol. Pharmaceutics* **2004**, *1*, 85-96.
10. Broccatelli, F.; Cruciani, G.; Benet, L.Z.; Oprea, T.I. Bddcs class prediction for new molecular entities. *Mol. Pharmaceutics* **2012**, *9*, 570-580.

11. Lindenberg, M.; Kopp, S.; Dressman, J.B. Classification of orally administered drugs on the world health organization model list of essential medicines according to the biopharmaceutics classification system. *Eur. J. Pharm. Biopharm.* **2004**, *58*, 265-278.
12. Takagi, T.; Ramachandran, S.; Bermejo, M.; Yamashita, S.; Yu, L.X.; Amidon, G.L. A provisional biopharmaceutical classification of the top 200 oral drug products in the united states, great britain, spain, and japan. *Mol. Pharmaceutics* **2006**, *3*, 631-643.
13. Wu, C.Y.; Benet, L.Z. Predicting drug disposition via application of bcs: Transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm. Res.* **2005**, *22*, 11-23.
14. Shawahna, R.; Rahman, N.U. Evaluation of the use of partition coefficients and molecular surface properties as predictors of drug absorption: A provisional biopharmaceutical classification of the list of national essential medicines of pakistan. *DARU* **2011**, *19*, 83-99.
15. Varma, M.V.; Gardner, I.; Steyn, S.J.; Nkansah, P.; Rotter, C.J.; Whitney-Pickett, C.; Zhang, H.; Di, L.; Cram, M.; Fenner, K.S., *et al.* Ph-dependent solubility and permeability criteria for provisional biopharmaceutics classification (bcs and bddcs) in early drug discovery. *Mol. Pharmaceutics* **2012**, *9*, 1199-1212.
16. Custodio, J.M.; Wu, C.Y.; Benet, L.Z. Predicting drug disposition, absorption/elimination/transporter interplay and the role of food on drug absorption. *Adv. Drug Deliv. Rev.* **2008**, *60*, 717-733.
17. Nair, A.K.; Anand, O.; Chun, N.; Conner, D.P.; Mehta, M.U.; Nhu, D.T.; Polli, J.E.; Yu, L.X.; Davit, B.M. Statistics on bcs classification of generic drug products approved between 2000 and 2011 in the USA. *AAPS J.* **2012**, *14*, 664-666.
18. *Volsurf+*, version 1.0.4; available from Molecular Discovery Ltd., London, U.K. (http://www.moldiscovery.com).
19. *Dragon for windows (software for molecular descriptor calculator).* 6.0; Talete srl, Milano Chemometrics and QSAR Research Group: http://www.talete.mi.it/products/dragon_description.htm.
20. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in the drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.
21. Chen, G.; Zheng, S.; Luo, X.; Shen, J.; Zhu, W.; Liu, H.; Gui, C.; Zhang, J.; Zheng, M.; Puah, C.M., *et al.* Focused combinatorial library design based on structural diversity, druglikeness and binding affinity score. *J. Comb. Chem.* **2005**, *7*, 398-406.
22. Ghose, A.K.; Viswanadhan, V.N.; Wendoloski, J.J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.* **1999**, *1*, 55-68.
23. Khandelwal, A.; Bahadduri, P.M.; Chang, C.; Polli, J.E.; Swaan, P.W.; Ekins, S. Computational models to assign biopharmaceutics drug disposition classification from molecular structure. *Pharm. Res.* **2007**, *24*, 2249-2262.
24. Zamora, I.; Oprea, T.I.; Ungell, A.L. Prediction of oral drug permeability. In *Rational approaches to drug design*, Holtje, H.D.; Sippl, W., Eds. Prous Science Press: Barcelona, Spain, 2001; pp 271-280.

25. Benet, L.Z.; Broccatelli, F.; Oprea, T.I. Bddcs applied to over 900 drugs. *AAPS J.* **2011**, *13*, 519-547.
26. *The international pharmacopoeia*. 4th ed.; World Health Organization: WHO Press, 20 Avenue Appia, 1211 Geneva 27, Switterland, 2006; Vol. 1 & 2, p 1520.
27. *The merck index*. 14th ed.; Merck Research Laboratories: Whitehouse Station, N.J., USA, 2006.
28. *Who model list of essential medicines* March 2011.
29. Electronic Orange Book. *Approved drug products with therapeutic equivalence evaluations*. 32nd ed.; Office of Generic Drugs Center for Drug Evaluation and Research, Food and Drug Administration: http://www.fda.gov/downloads/Drugs/InformationOnDrugs/UCM086233.pdf, updated August 2012.
30. Pham The, H.; Gonzalez Diaz, I.; Bermejo Sanz, M.; Mangas Sanjuan, V.; Centelles, I.; Garriges, T.M.; Cabrera Perez, M.A. In silico prediction of caco-2 permeability by a classification qsar approach. *Mol. Inf.* **2011**, *30*, 376-385.
31. Kim, J.S.; Mitchell, S.; Kijek, P.; Tsume, Y.; Hilfinger, J.; Amidon, G.L. The suitability of an in situ perfusion model for permeability determinations: Utility for bcs class i biowaiver requests. *Mol. Pharmaceutics* **2006**, *3*, 686-694.
32. Netzeva, T.I.; Worth, A.P.; Aldenberg, T.; Benigni, R.; Cronin, M.; Gramatica, P.; Jaworska, J.S.; Kahn, S.; Klopman, G.; Marchant, C.A., *et al.* Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ecvam workshop 52. *ATLA* **2005**, *33*, 1-19.