ISSN: 0975-766X
CODEN: IJPTFI
Available through Online ***Review Article***
www.ijptonline.com

IN SILICO QUANTITATIVE STRUCTURE PHARMACOKINETIC RELATIONSHIP MODELING ON ANTIDIABETIC DRUGS: APPARENT VOLUME OF DISTRIBUTION

Sethi Reeta[1] and Paul Yash[2]
[1]Shri Jagdish Prasad Jhabarmal Tibrewala University, Jhunjhunu (Rajasthan), India.
[2]Lord Shiva College of Pharmacy, Sirsa (Haryana), India.
Email: reetasethi05@gmail.com

Received on 28-02-2016 *Accepted on 27-03-2016*

Abstract

The use of *in silico* approaches for successful prediction of pharmacokinetic properties of compounds during new drug discovery has been increasing exponentially. These *in silico* models, for the prognosis of absorption, distribution metabolism and excretion (ADME), are invariably based on the implementation of quantitative structure pharmacokinetic relationship (QSPkR) techniques. The current study was conducted to investigate QSPkR for apparent volume of distribution (Vd) in man among 24 Antidiabetic drugs employing an extra thermodynamic approach. It is vital to predict the Vd value of various drug leads during drug discovery so that compounds with poor bioavailability can be eliminated and those with an acceptable metabolic stability can be identified. Analysis of several thousands of QSPkR correlations developed in the present study revealed an extremely high degree of cross-validated coefficient (Q^2) using the leave-one-out method ($P < 0.001$). Logarithmic transformation tends to improve the correlations marginally ($R^2 = 0.9347$) but the inverse transform resulted in a distinct improvement in the correlation ($R^2 = 0.9932$). Electronic and topological parameters were found to primarily ascribe the variation in Vd. Overall, the diffusional interactions seem to play a major role in attributing V_d rather than the permeational ones.

Key words: ADME prediction, pharmacokinetics, quantitative structure pharmacokinetic relationship, QSPkR

Introduction

It is now duly recognized by the pharmaceutical industry that undesirable absorption, distribution, metabolism and excretion (ADME) of new drug candidates are the cause(s) of many clinical phase drug development failures. Nearly 45% of the drug candidates fail during the clinical trials owing to poor pharmacokinetic properties. This is an economic disaster

as the failed drugs have been in the pipeline for several years with a huge expenditure of efforts, time and money invested in their development. Accordingly, it has been an earnest endeavor of the pharmaceutical scientists to identify such problems early during the drug delivery process and design new drug molecules with optimal pharmacokinetic and pharmacodynamic properties before their synthesis.

Of late, the in vitro approaches have been widely practiced to investigate the ADME properties of new chemical entities.[1] More recently, *in silico* modeling has been investigated as a tool to optimize selection of the most suitable drug candidates for development.

This novel approach of quickly predicting the ADME properties using computational means is of great importance because the experimental ADME testing is phenomenally expensive and arduous. Therefore, the use of computational models in the prediction of ADME parameters has been growing rapidly in drug discovery because of their immense benefits in throughput and early application of drug design.[2]

Apparent volume of distribution (V_d), a vital pharmacokinetic parameter characterizing the dispositional attributes of a drug, is simplistically a proportionality constant relating the plasma drug concentrations to the total amount of drug in the body. [3] Its magnitude provides a broad inkling as to how widely the drug gets distributed in the body. Also, V_d provides an excellent tool to correlate the physicochemical properties with the duration and intensity of action based on its distribution in the body. Depending on the degree of drug binding to the plasma proteins and tissues, large variations are noticeable in the apparent volume of distribution of various drugs in man. An estimate of V_d is of paramount importance while selecting a drug candidate in therapeutics and while calculating its biological half life and clearance values.[4]Traditionally, the V_d value of a drug candidate is obtained via in vivo studies, which tends to be quite arduous, time consuming and expensive. The *in silico* ADME modeling using the quantitative structure pharmacokinetic relationship (QSPkR) method has been explored for predicting the V_d value of drug candidates in an efficient and cost-efficient manner.[5-8]. The primary aim of these QSPkR studies is to enable the drug designer to modify the chemical structure of a pharmacodynamically active drug in such a manner as to alter its pharmacokinetic properties without diminishing its pharmacodynamic potential.[9-10]

The major advantage of QSPkR, therefore, lies in the fact that once such a relationship is ascertained with an adequate statistical degree of confidence, it can be a valuable assistance in the prognosis of the behavior of new molecules even

before they are actually synthesized.[11]The key objective of the current study was to investigate *in silico* QSPkR among various Antidiabetic drugs for V_d. Antidiabetics were chosen for QSPkR studies as this category of drugs is extensively used as in the treatment of diabetes. Also, Antidiabetics consist of a significant number of drug compounds thoroughly investigated for their pharmacokinetic performance, particularly for V_d (n = 24).Further, the congeners in this class have many common pharmacokinetic characteristics, mechanism and degree of affinity with body tissues. Several descriptors like experimental values of log P, pKa, melting point, etc. of these drugs are available in the standard texts or journals.

Materials and methods

QSPkR was conducted among Antidiabetic drugs employing an extra thermodynamic multilinear regression analysis (MLRA) approach. The general steps for developing a QSPkR model include data set selection, chemical structure entry, 3D structure generation and descriptor calculation, model construction that involves selection of descriptors and validation of the testing set using a Pentium dual core microprocessor (Intel, Santa Clara, USA) desktop (IBM, Bangalore, India) with 1 GB RAM and 160 GB hard diskdrive. The computer peripherals included an HP Laser 1020 series printer and an HP Scanjet 2400 scanner.

Dataset selection

The reported values of V_d of the Antidiabetic drugs in humans were taken from various literature sources.[3,12-15] In order to ensure that experimental variations in determining V_d do not significantly affect the quality of our datasets, only V_d values obtained from healthy adult males after oral administration were employed for constructing the dataset. A total of 24 Antidiabetic drugs were selected and used as the dataset for this study.

The V_d value of each of these compounds was also log-transformed (log V_d) and inverse transformed (1/ V_d) to normalize the data and to reduce the unequal error variance, respectively.

Molecular structure and descriptors

Various structural parameters were computed theoretically employing diverse computer software.

Descriptors calculated by Pallas 2.0

The values of structural descriptors, like log P, pK_a and log D of the various Antidiabetics, were calculated using the software Pallas 2.0 (CompuDrug International Inc., Sedona, USA). The structures of the drugs were graphically drawn on

the monitor with the help of a mouse. Suitable templates/rings were chosen, bonds were drawn and different hetero atoms were chosen from the periodic table provided in the software and incorporated into the structure. The rough graphical sketch representing the structural formula of the compound was transformed to its least-energy configuration. The name of the compound was entered to let the structure of the drug be stored under its assigned name in the software database. For the estimation of log P and log D, compounds from the database were selected, the software run for estimation of the desired descriptors and the results were stored as an MDL molfile.

Descriptors calculated by HyperChem

Log P, pKa, surface area and surface volume of various Antidiabetics were calculated using the software HyperChem 8.0.5 (Hypercube Inc., Gainesville, USA). The structures of the drugs were graphically drawn on the monitor and the same procedure was followed as described above in descriptors calculated by Pallas 2.0 and, however, at the end in this case hin files were generated.

Parameters calculated by Dragon

The molfiles generated by the Chem 3D software pro v.3.5. (Cambridge Soft Corporation, Cambridge, MA, USA) were imported to Dragon 5.5 (Talete Srl, Milano, Italy). As many as 1497 diverse descriptors, viz constitutional, geometrical, topological, Whim 3D, electronic, etc., were calculated with the help of the Dragon software.

Parameters calculated by CODESSA

A large number of molecular descriptors were calculated with the help of the CODESSA 2.0 software (Semichem, Shawnee, Terrace, USA) also. First of all, a worksheet was made in an MS- Excel environment to load various molfiles into the software. The file was saved as a nondocument ASCII text file. The said text file consisted of a number of columns separated by blanks, each column containing data of one type, e.g. structure names, property values, file names, etc. Each line contained the same number of columns. The program then scans the file in order to determine the number of columns and provides a column dialog box, where the type of data in each column and other parameters were specified. Before calculating the descriptors, the loaded structure was checked and necessary corrections were made. A "structure dialog box" was used to enter or change the structure name as well as names and type of files associated with the structure. Various classes of descriptors, viz constitutional, topological, geometrical and electronic descriptors, were selected for calculation using the "calculate descriptor" dialog box. Initially, the descriptors were computed for all the structures

loaded into the software. Further, as and when any information was available about new congeners, those particular compounds were also selected for computation of the descriptors.

Multivariate statistical analysis

Attempts were made to correlate all the types of descriptors, viz lipophilic, constitutional, electrostatic, electronic, topological and steric, with the pharmacokinetic parameter V_d. The initial regression analysis was carried out using a heuristic analysis followed by the best multilinear regression and MLRA options of the CODESSA software. In case of the heuristic method, a pre selection of the descriptors was accomplished. All the descriptors we rechecked to ensure that the value of each descriptor was available for each structure with significant variation among these values. Descriptors, for which values were not available for every loaded structure in the data, were discarded. Thereafter, the one-parameter correlation equations for each descriptor were calculated. The number of descriptors in the starting set was further reduced by discarding if:

• The F-value for the one-parameter correlation with the descriptor is below 1.0.

• The r^2 value of the one-parameter equation is less than the assigned value of r^2 min (usually 0.1).

• The one-parameter *t*-value is less than the assigned value (usually 1.5).

• The multi parameter *t*-value is less than the assigned value (usually 1.93).

• The descriptors are highly inter correlated with another descriptor ($r^2 > 0.66$).

Pharmacokinetic data of the V_d parameter, available for 24 Antidiabetics, were analyzed limiting the descriptors: drug ratio to 1:4. The heuristic method yielded a list of the best 10 correlations, each with the highest values of R^2 and F ratio. Numerous attempts were carried out to obtain significant correlations for Antidiabetics, some of which are shown in Table A.

Equations	**M**	**R^2**	**F**	**S^2**	**Q^2**	**p<**
V_d= -3.1614 + 0.7088 ACIC1	1	0.3153	11.13	0.6921	0.2426	0.005
V_d= 1.0206 + 0.2713 AIC0 – 41.0215 QNmax	2	0.4834	12.69	0.4468	0.4009	0.001
V_d= 3.14 + 0.4115 ABIC1 – 33.4569 QNmax + 0.0974 Es	3	0.5756	15.84	0.3700	0.5736	0.001
V_d= -5.2124 - 0.2157 VDW + 0.6024 ABIC1 - 48.0214 QNmax - 9.465 ZXS/ZXR - 6.0151 MSA	5	0.7926	26.05	0.2406	0.7608	0.001

Equations	M	R^2	F	S^2	Q^2	p<
V_d= -1.003 + 4.2415 ASIC0 + 0.231 ABIC2 + 0.0181 Es - 0.3125 VDW + 71.9001 Qmin + 1.4217 XYS	6	0.9126	29.26	0.1139	0.8601	0.001
Log V_d = -0.6237 – 0.1164 ACIC1	1	0.3751	13.94	0.0303	0.3128	0.001
Log V_d = -4.0052+ 0.9362 MSA – 3.5346 ASIC0	2	0.5661	16.10	0.0210	0.4836	0.001
Log V_d = 1.4321 + 0.3382 AIC0 + 0.0017 ASIC2 - 6.0017 QNmax	3	0.6847	25.36	0.0174	0.5613	0.001
Log V_d = 4.3931 - 5.3612 QNmax + 0.3741 VDW + 11.321 E3s -0.012311 ZXS	4	0.7909	30.82	0.0103	0.6983	0.001
Log V_d = 2.3650 + 0.0163 AIC1 + 0.0031 G3v - 0.0015 Es -114.45 On + 0.0599 Tm	5	0.8807	38.69	0.0077	0.8109	0.001
Log V_d = 1.0755 - 9.7642 ABIC1 + 12.104 Qmin - 0.00621 E3e + 0.25432 VDW - 7.1310 QNmax – 3.5041 ASIC0	6	0.9347	40.16	0.0047	0.8913	0.001
1/V_d= -0.4323 + 24.400 G3m	1	0.6238	56.37	0.2804	0.3811	0.001
1/V_d = 11.36 + 86.411 QNmin - 10.469 AIC0	2	0.7506	63.57	0.1632	0.5922	0.001
1/V_d = 8.712 - 7.1422 YZS/YZR -0.24321 QNmax – 12.678 ASIC1	3	0.8226	74.52	0.1083	0.7176	0.001
1/V_d = -0.0612 - 0.1153 VDW - 0.0212 ZXS + 12.621 Tm	4	0.8936	103.44	0.0724	0.7884	0.001
1/V_d= –5.4321 + 6.0323 ACIC1 -0.3224 AIC0 + 83.2141 QNmin + 18.934 G3m + 0.0204 WPSA-2	5	0.9619	141.86	0.0382	0.8900	0.001
1/V_d=11.836 + 9.9123 MSA + 85.432 QNmin - 10.693 ASIC0 + 12.3174 G3m + 3.9213 E3s - 0.0020 PNSA-2	6	0.9932	178.35	0.0204	0.9370	0.001

Table A: Significant linear, logarithmic and inverse QSPkR equations for a series of 24 Antidiabetic drugs using V_d as pharmacokinetic parameter.

A set of important descriptors found to significantly ascribe the variation of V_d was constructed. Further, a search for the Multi parameter regression with the maximum predicting ability was performed. A number of sets of descriptors were thus made and MLRA was performed with V_d. Regression plots of each correlation thus attempted were examined for linearity and coherence. Residual plots were also examined for randomization and absence of distinct patterns in order to

eliminate chance correlations [Figure 1]. Logarithmic and inverse transformations of V_d were also carried out in order to screen the correlation with improved values of R^2 and/or F ratio [Figures 2 and 3]. Graphs were constructed using the MS-Excel software.

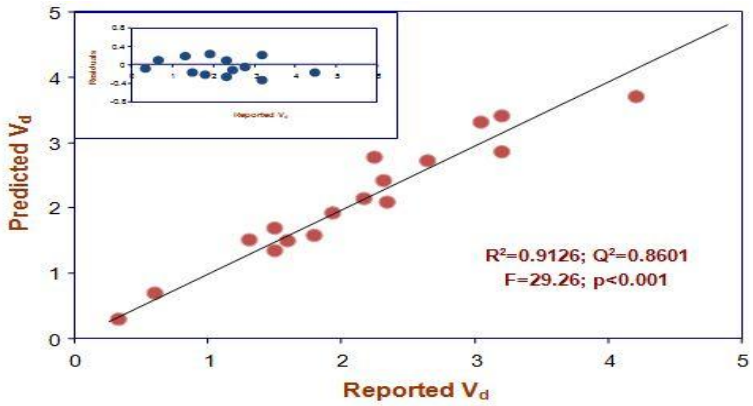

Figure 1: Linear correlation plot between the values of V_d as reported in the literature and those predicted using the multi parameter quantitative structure pharmacokinetic relationship for a series of 24 Antidiabetics. The inset shows the corresponding residual plot.

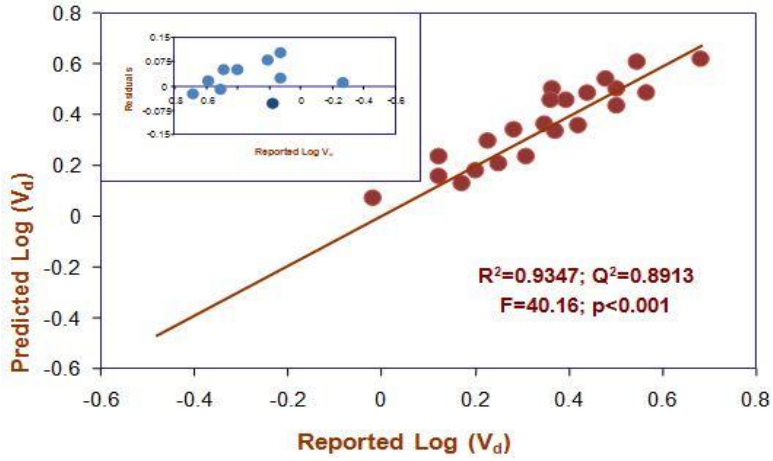

Figure 2: Linear correlation plot between the values of log transform of V_d as reported in the literature and those predicted using the Multi parameter quantitative structure pharmacokinetic relationship for a series of 24 Antidiabetics. The inset shows the corresponding residual plot.

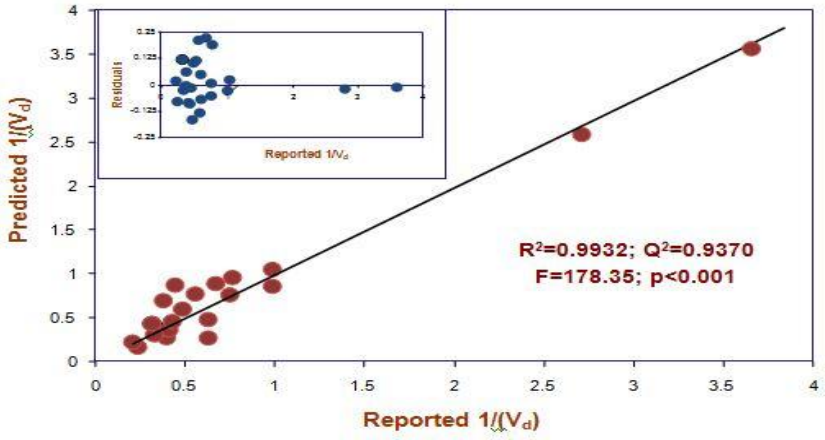

Figure 3: Linear correlation plot between the values of inverse transform of V_d as reported in the literature and those predicted using the multi parameter quantitative structure pharmacokinetic relationship for a series of 24 Antidiabetics. The inset shows the corresponding residual plot.

Validation of the testing set

Statistical significance of each correlation was determined on the basis of the value of the F-criterion and the magnitude of the cross-validated R^2, commonly represented as Q^2, calculated according to Equation no.1.

$$Q^2 = 1 - \frac{\sum (ypred - yobs)^2}{\sum (yobs - ymean)^2} \qquad \ldots (1)$$

A model with good predictive performance will have a Q^2- value close to 1, models that do not predict better than merely chance alone can have negative values.

The F-values were computed according to Equation no. 2:

$$F = \frac{S_1^2}{S_2^2} \qquad \ldots (2)$$

Where, *S*1 is the variance between the samples and *S*2 is the variance within the samples.The values of the computed F-ratio were compared with that of the critical values tabulated in the statistical texts and the levels of significance were discerned. The QSPkR correlations found to be statistically significant were compiled from the CODESSA software and were stored as respective files under the extension of COD. The names of descriptors were conveniently coded using a WS-Macro program and the files were converted to an appropriate ASCII format using in-house developed program codes. These ASCII files were further converted into tabular formats in MS-Word.

Results and discussion

Variable QSPkR results were obtained following the application of multivariate statistical analyses on Antidiabetic drugs. Thousands of such correlation and regression analysis were attempted choosing all the possible combinations of available descriptors, each yielding an elaborate output. The concise results of only those correlations that were found to be statistically significant, usually at a 5% level or less, and/or those that have important applications have been taken into consideration. The volume of distribution for a combined set of 24 Antidiabetics showed significant dependence on the topological parameters and geometric parameters. The prominent descriptors explaining variation in V_d encompass

the information contents, structured information contents, hydrophilic factor (Hy) and other parameters like shape profile no. 02 (SP02), 3st component symmetry directional WHIM index/ weighted by atomic Sanderson electro negativities (G3e), 3st component symmetry directional WHIM index/ weighted by atomic polarizabilities (G3p), number of H atoms (Hn) and folding degree index (FDI). The electronic parameters like Max partial charge for N atom (Zefirov's PC) (QNmax) and Min partial charge for N atom (Zefirov's PC) (Qnmin) and geometrical parameters like XY Shadow (XYS) also yielded minor contributions toward improvement in relationships. Thus, overall, the diffusional interactions seem to play a pivotal role in attributing V_d rather than the permeational ones.

Logarithmic transformation tends to improve the correlations marginally ($R^2 = 0.9347$) but inverse transforms resulted in a distinct improvement in the correlation ($R^2 = 0.9932$). Dependence on the nature of the descriptors remained similar for log-transformed values. However, for the inverse transformed values, increased dependence on WHIM descriptors (like G3e, G3P, G3m) was noted. Earlier studies[16-18] have correlated volume of distribution to lipophilicity. Our results in the current studies, on the contrary, show dependence of V_d more on topological and electronic parameters than on lipophilic parameters. It can be very well explained on the basis of the involvement of ionic bonding and van der Waal's interactions that play a major role in tissue and protein binding thus affecting the V_d.[19] The primary reason for the difference in the outcomes might be the involvement of numerous descriptors of a varied nature in our study vis-à-vis only a limited number of mainly lipophilic descriptors involved in the earlier reports.

Conclusions

Highly significant results on *in silico* prognosis of V_d ($P < 0.001$) attributed major variation to the electronic and topological descriptor, vouching the dependence on the diffusional interactions. Chance correlations, if any, were ruled out in the light of high magnitudes of cross-validated variance, i.e. Q2, obtained in the current QSPkR studies. Pharmacokinetic performance of a drug is known to be not merely a function of its physicochemical nature but of the biological system(s) too, like somatic, psychological, environmental, nutritional, genetic, hereditary and diurnal status of the human subjects. This causes a great deal of plausible variation in pharmacokinetic profiles among the volunteers/patients undergoing the study. The literature values of the pharmacokinetic parameters taken up in the present investigations pertain to diverse subject populations hailing from different age groups, genders, races, nutritional and physical attributes, etc. studied in different geographical regions under different weather conditions. Considering these

potentially high inter subject and intra subject variations among the pharmacokinetic parameters, the currently established relationships assume much higher credibility. It seems highly probable that the *in silico* approaches will evolve rapidly, as did the *in vitro* methods during the last decade. Past experience with the latter could be helpful in avoiding repetition of similar errors and in taking the necessary steps to ensure effective implementation of the former.

References

1. Waterbeemed VH, Gifford E. ADMET *in silico* modeling: Towards prediction paradise? Nature Reviews. Drug Discov 2003;2:192-204.

2. Beresford AP, Segall M, Tabit MH. *In-silico* prediction of ADME properties: Are we making progress? Curr Opin Drug Discov Develop 2004;7:36-42.

3. Shargel L, WU-Pong S, Andrew YU. Applied Biopharmaceutics and Pharmacokinetics. 5th ed. McGraw Hill Companies Inc; 2005. p. 259,8646.

4. Ghafourian T, Barzegar JM, Dastmalchi S, Khavari-Khoraseni T, Kakimiha N, Nokhodchi A. QSPkR models for the prediction of apparent volume of distribution. Int J Pharm 2006;319:82-97.

5. Karlis V, Tsantili-Kakoulidou A, Macheras P. Quantitative structure pharmaockinetic relationships for disposition parameters of cephalosporins. Eur J Pharm Sci 2003;20:115-23.

6. Turner JV, Maddalena DJ, Cutler DJ, Agatonovic-Kustrin S. Multiple pharmacokinetic parameter prediction for a series of cephalosporins. J Pharm Sci 2003;92:552-9.

7. Saha RN, Raman N. Modeling of Biological Activity and Pharmacokinetics of Cephalosporin Derivatives. Ind J Pharm Edu Res 2008;42:207-14.

8. Ng C. Xiao YD, Putnam W, Lum B, Tropsha A. Quantitative structure pharmacokinetic parameters relationships (QSPkR) analysis of antimicrobial agents in humans using simulated annealing k-nearest- neighbor and partial least-square analysis methods. J Pharm Sci 2004;93:2535-44.

9. Seydel JK. Structure pharmacokinetic relationship in drug design. Amsterdam: Elsevier; 1982. p. 179.

10. Grover M, Singh B, Bakshi M, Singh S. Quantitative structure property relationships in pharmaceutical research - Part-2. Pharm Sci Technol Today 2000;3:50-7.

11. Singh B, Dhake AS, Sethi D, Paul Y. *In silico* ADME predictions using Quantitative structure Pharmacokinetics Relationships Part-I: Fundamental Aspects. Pharm Rev 2007;5:93-100.

12. Hooper DC, Wolfson JS. Quinolones Antimicrobial Agents. 2nd ed. Washington: American Society for Microbiology; 1993. p. 195-223.

13. Mignot A, Guillaume M, Gohler K, Stahlberg HJ. Oral Bioavailability of gatifloxacin in healthy volunteers under fasting and fed conditions. Chemotherapy 2002;48:111-5.

14. The Merck Index: An Encyclopedia of Chemicals, Drugs and Biological. 4th ed. USA: Merck Research Laboratories Division of Merck and Co INC; 2006.

15. Bruton LL, Laza JS, Parker KL. Goodman Gilman's: The Pharmacological Basis of therapeutics. 11th ed. New York: McGraw-Hill Medical Publishing Division; 2006. p. 940-50.

16. Watari N, Sugiyama Y, Kaneniwa N, Hiura M. Prediction of hepatic first-pass metabolism and plasma levels following intravenous and oral administration of barbiturates in the rabbit based on quantitative structure-pharmacokinetic relationships. J Pharmacokinet Biopharm 1988;16:279-301.

17. Hirono S, Nakagome I, Hirono H, Matsushita Y, Yoshi F, Moriguchi I. Non-congeneric structure-pharmacokinetic property correlation studies using fuzzy adaptive least-square, oral bioavailability. Biol Pharm Bull 1994;17:308.

18. Blakey GE, Nestorov IA, Arundel PA, Aorons LJ, Rowland M. Quantitative structure-pharmacokinetics relationships (QSPkR), 1. Development of a whole-body physiologically based model to characterize changes in pharmacokinetics across a homologous series of barbiturates in the rat. J Pharmacokinet Biopharm 1997;25:277.

19. Houglum J, Harrelson G, Leaver-Dunn D. Principles of Pharmacology for Athletic Trainers. USA: Slack Incorporated; 2005. p. 19-82.

Corresponding Author:

Reeta Sethi*,

Email: *reetasethi05@gmail.com*