INTERNATIONAL ASSOCIATION OF BIOLOGICALS AND COMPUTATIONAL DIGEST
International & Peer-Reviewed Journal
E-ISSN: 2583-3995

QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIP (QSAR) STUDY OF LIVER TOXIC DRUGS

[1]Nandan Dixit, [1]Chirag Patel, [1]Mansi Bhavsar, [1]Saumya K Patel, [2]Rakesh M Rawal and [1]Hitesh A. Solanki*

[1]Department of Botany, Bioinformatics and Climate Change Impacts Management, Gujarat University, Ahmedabad-380009, Gujarat, India.
[2]Department of Life Science, School of Sciences, Gujarat University, Ahmedabad-380 009.

Email: husolanki@yahoo.com

ABSTRACT

Drug-induced liver injury (DILI) is one of the most severe adverse effects (AEs) causing life-threatening conditions, such as acute liver failure. t has also been recognized as the single most common cause of safety- related post-market withdrawals or warnings Due to the nature and idiosyncrasy of clinical forms of DILI, attempts to develop new predictive approaches to evaluate the risk of a medication being a hepatotoxicant have been difficult. The FDA Adverse Event Reporting System (AERS) provides post-market data illustrating AE morbidity. A quantitative structure –activity relationship (QSAR) model for DILI prediction with satisfactory output is urgently needed. In this study, we documented a high-quality QSAR model for predicting the hepatotoxicity risk of DILI by integrating the use of eight effective classifiers and molecular descriptors given by the VlifeMds program. For the present QSAR study, data set of 99 compounds (withdrawn and approved drugs) collected from different databases were taken. Multiple linear regression and partial least square analysis methods had developed two dimensional QSAR models, and then validated for internal and external predictions. The 2D QSAR model developed was statistically important, and was highly predictive. The validation methods presented essential statistical parameters that proved the model's predictive ability. The developed 2D QSAR model revealed the significance of SsssCE-index, SsOHcount, SsssNcount and SdssPcount descriptors. These findings will prove to be an important guide for furtherdesigning and developing new hepatotoxicity activity.

Keywords: *Hepatotoxicity, QSAR, DILI*

INTRODUCTION

A long and arduous method is the discovery and creation of a new chemical agent with proven usefulness in ameliorating or curing disease.According to industry estimates, up to several thousand compounds are synthesized and evaluated; up to 100 compounds are assessed forsafety; and up to 10 compounds are clinically tested in humans for each medicinal product approved.Trial and error screening, used as the standard method, is becoming very costly and less effective at the same time.Therefore only molecules should be prepared and tested with goodchances of action.Proper design is necessary in this context before synthesizing the drugs[3].The liver plays a critical role in energy exchanges as the first organ that comes in contact with mostof the digestive products.Damaged liver often disrupts normal metabolism and even contributes to liver failure[1].Over the last decades, drug-induced liver injury (DILI) continues to be an important area of study as one of the main types of liver damage.DILI has been described as the most frequent reason for stopping drug research projects in the drug discovery process.In addition, over the past half century hundreds of medicines were removed from the market and refused for evidence of liver damage in the latest drug applications[2].Eliminating drug candidates with DILI risks early in the drug discovery may be an effective strategy for reducing the attrition rate and reducing drug discovery costs.Therefore more focus should be paid to work that aims to determine the DILI risk of

drugs and drug candidates.The hepatotoxicity of the drugs has historically been experimentally observed.Yet one can not ignore that certain forms of experimentation are time-consuming and labor-intensive.Furthermore, DILI induced by most drugs is of an idiosyncratic nature and can not typically be detected by the regulatory experiments needed for animal / cell toxicity[8].Compared to experimentally detecting hepatotoxicity, predicting the risk of DILI in silico models is more time-saving and low-cost and is effective in assessing drug candidates' potential DILI risk[11,12].
Many 2D QSAR studies have been reported for different group of chemicals derivatives in rational drug design.With the aim of developing good drugs, we have selected withdrawal drugs that make DILI and Approved compounds understand structural insight, which is responsible for the selectivity of these drugs towards hepatotoxicity using QSARanalysis[9,10].The series of compounds had demonstrated well defined activity.The series of drugs selected for the present study contained high structural diversity and a sufficient range of biological activity.The QSAR models developed have been statistically significant and could effectively guide the preparation of potential algorithms to discriminate against acceptable vs. toxic drugs.

MATERIALS AND METHODS

1.1 Selection of molecules
The data set for the present QSAR study was taken from 99 compounds collected from different databases.Of these 99 compounds, 43 compounds were withdrawn from the market because of the cause of hepatotoxicity, while 56 compounds were approved as [supplementary material] drugs.
To get the linear relationship in equation, the Lethal dose (50) activity data of the compounds were converted into [-Log10(mol / kg)]].
Based on suggestions, molecules were widely separated into the training set and the test set.

1.2 Molecular modelling
All computational experiments were carried out using the Molecular Design Suite (VlifeMDS) software on HP computers with incontrovertible Intel Pentium Dual Core Processor and Windows XP operating system. Structures were drained using the 2D draw application and subjected to an energy minimization and geometry optimization using Merck Molecular Force Field, force field and charges followed by Austin Model-1 with 10000 as maximum number of cycles, 0.01 as meeting criteria *(*root mean square gradient *)*and 1.0 as constant *(*medium's dielectric constant which is 1 for in vacuo *)*in dielectric properties .The default values of 30.0 and 10.0 Kcal/mol were used for electrostatic and steric energy cut off[13,15].

2.32D-QSAR analysis

2.3.1 **Calculation of descriptors** figures of descriptors were calculate after optimization or minimization of the energy of the data set molecules. a range of types of physicochemical descriptors were calculated [15].

2.3.2 Generation of training and test sets:

In command to evaluate the QSAR model, data set was separated into training and test set using sphere exclusion, random selection and manual selection method. Data set of 99 compounds collected from various database were taken for the present QSAR study . Among these 99 compounds 43 compounds were selected for test set while 56 compounds were selected for training set size [13,15].
Training set is used to develop the QSAR model for which biological activity data are known. Test set is used to challenge the QSAR model developed based on the training set to assess the predictive power of the model which is not included in model generation.
Sphere Exclusion method: In this process at first data set were alienated into training and test set using sphere exclusion method .In this method contrast value provides a design to handle training and test set size .It needs to be accustomed by trial and error until a required division of training and test set is achieved .Increase in contrast value results

in increasein number of molecules in the test set.

Random Selection Method: In order tobuild and validate the QSAR models, in cooperation internally and externally, the data sets were separated into training and test in a arbitraryway .

Manual data selection method: Data set is alienated manually into training and test sets on the basis of the outcome obtained in sphere exclusion method and random selection method.

2.3.3 Generation of 2D-QSAR models:

Two dimensional quantitative structure activity relationship (2D QSAR)studies by way of multiple linear regression (MLR)method was perform on a seriesData set of 99 compounds (approved & withdrawn)of via software QSAR (VLife Science.)MLR is a technique used for modelling linear relationship between a dependent variable Y (Activity)and independent variable X (2D descriptors .)MLR is based on smallest amount squares .The model is robust such that sum-of squares of differences of observed and a predicted value is minimized . MLR calculate approximately values of regression coefficients (r2)by apply least squares curve appropriatemeans .The model creates a relationship in the form ofa straight line (linear)that finest approximates all the entity data points .In regression analysis, conditional mean of dependant variable (Activity)Y depends on (descriptors)X .MLR analysis extends this idea to include more than one independent variable.

RESULT AND DISCUSSION

Different sets of 2D-QSAR models were generated using the MLR analysis in combination with stepwise forward-backward variable selection method. Different training and test set were constructed using sphere exclusion, random and manual selection Method. Training and test set were selected if they follow theunicolumn statistics, i.e. maximum of the test is lessthan maximum of training set and minimum of the test set is greater than of training set, which is prerequisite for further QSAR analysis. This result shows that the test is interpolative i.e., derived from the min-max range of training set.

The mean and standard deviation of the training and test set provides insight to the relative difference of mean and point density distribution of the two sets. The statistical significant 2D-QSARmodels for column “log p activity distribution.” The selection of the best model is based on the values of r2(squared correlation coefficient), q2(cross-validated correlation coefficient), pred_r2(predicted correlation coefficient for the external test set),F (Fisher ratio) reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F–test indicate that the model is statistically significant. r2se, q2se and pred_r2se are the standard errors terms for r2, q2and pred_r2respectively.

In the above QSAR equation, n is the number of molecules (Training set) used to get the QSAR model, r2is the squared correlation coefficient, q2is the cross-validated correlation coefficient, pred_r2is the predicted correlation coefficient for the external test set, F is the Fisher ratio, reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant. r2se, q2se and pred_r2se are the standard errors terms for r2, q2and pred_r2 (smaller is better).

1.3 2D-QSAR model explains n = 56, r 2 = 0.658, r2 se = 0.327, q2 = 0.530, q2 se = 0.383, F_Test = 19.270, pred_r 2 = 0.233, pred_r2 se = 0.4471, Z Score r 2 = 13.156, Z Score q2 = 5.384, Best Rand r2 = 0.243, Best Rand q2 = 0.103 65.84 % (r2= 0.6584) of the total difference in the training set as well as it has internal (q2) and external (pred_r2) predictive ability of 44.71 %. . From QSAR model, SssssCE-index: Electrotopological state indices for number of carbon atom connected with four single bonds.SsOHcount: This descriptor defines the total number of –OH group connected withone single bond.SsssNcount: This descriptor defines the total number of nitrogen connected with three single bonds.SdsssPcount: This descriptor defines the total number of phosphorous atom connected with three single bonds and one double bond.

Contribution chart for model is represented in Figure 1 reveals that the descriptors SSssNcount and DEltaEpsilonB contributing approximately 15.5 %, 14.5 % respectively.

Three more descriptors SssssCEindex ,SsOHcount and SdsssPcount are contributing inverselyapproximately 33 %, 19 %, and 17% respectively to activity.

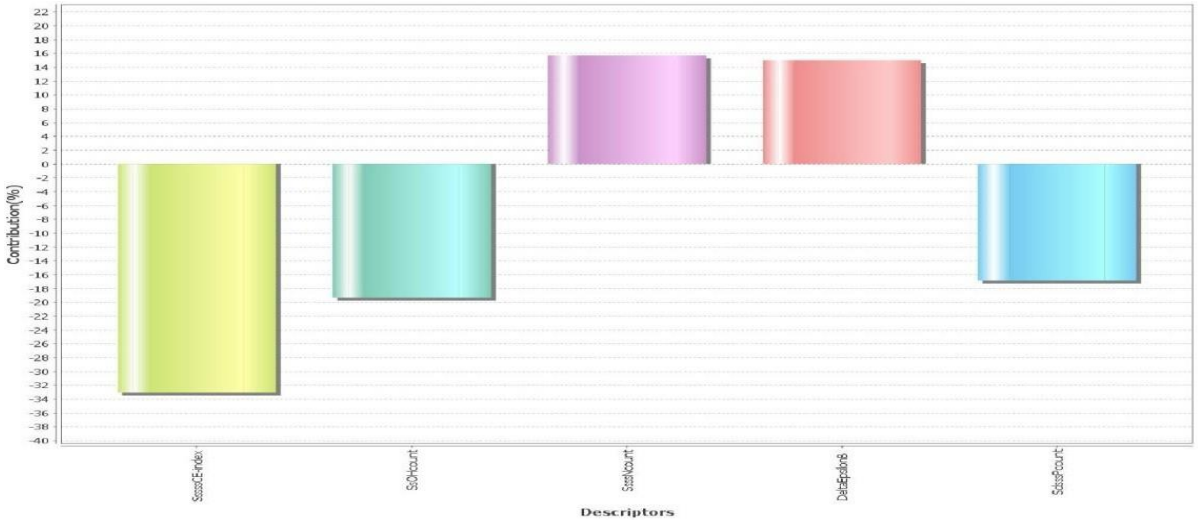

Figure 1 Contribution chart for 2D QSAR model showing contribution of different descriptors

Data fitness plot for 2D- QSAR model is shown in Figure 2. The plot of observedvs predicted activity provides an idea about how well the model was trained and how well it predicts the activity of external test set

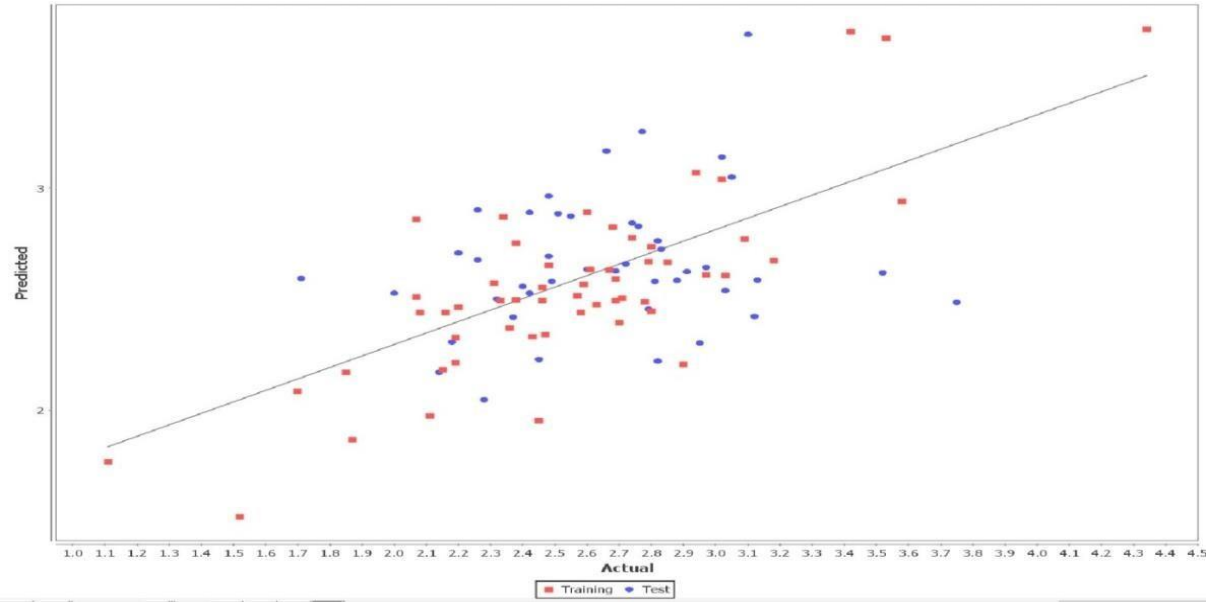

Figure 2 Data fitness plot for 2D QSAR model

The graph of observed vs. predicted activity of training and test sets for 2D-QSAR modelis shown in Figure 3 .the2D-QSAR model is able to predict the activity of training set quite well as well as external test set, providing confidence of model. Results of the observed and predicted activity are shown below.
Test Set and Training Set (Actual Vs Predicted)

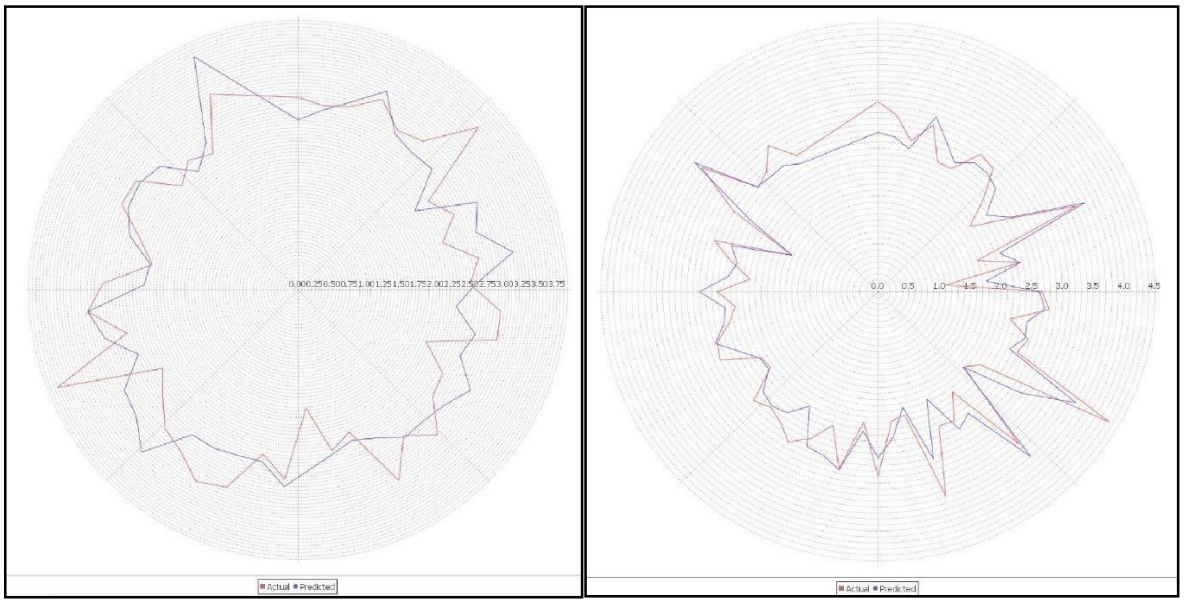

Figure 3 Radar plot Actual vs predicted test set and training set of 2D QSAR

CONCLUSION

In the present study, statistically significant and highly predictive 2D QSAR models were developed for some Liver toxic compounds. The QSAR models were validated by standard statistical measures, crossvalidated correlation coefficient, external test set and randomization test, and through observation on how it reproduces and explains the quantitative differences seen in the experimentally known activity data. The models are considered predictive model as the validation methods provided significant statistical .The developed 2D QSAR models revealed the importance of SssssCE-index,SsOHcount,SsssNcountand SdsssPcount:properties of compounds . These results will be an essential guide for the further design and development of new lead compounds of more potent antidiabetic compounds. These studies can be further extended to develop QSAR models using some other approaches 3D QSAR and docking analysis of direct drug designing and further validation of the results obtained in the present studies. The field is further open for designing, synthesis and biological evaluation of potent antidiabetic compounds, pharmacokinetic studies and clinical studies to establish those molecules asdrug.

REFERENCES

1) http://cheminfo.charite.de/withdrawn/
2) https://www.drugbank.ca/
3) Benton, C. B., Nazha, A., Pemmaraju, N., Garcia-manero, G., Disorders, B., Cancer, T., & Clinic, C. (2016). HHS Public Access, *95*(2), 222–242. https://doi.org/10.1016/j.critrevonc.2015.03.002.Chronic
4) Taniguchi, C. M., Armstrong, S. R., Green, L. C., Golan, D. E., &Tashjian, A. H. (2007). Mechanisms of Drug Toxicity. *Drug Toxicity*, 63–74.
5) Farghali, H., Kemelo, M. K., &Wojnarová, L. (2016). In Vitro and In Vivo Experimental Hepatotoxic Models in Liver Research : Applications to the Assessment of Potential Hepatoprotective Drugs, *65*.
6) Jaeschke, H., & City, K. (2016). HHS Public Access, *33*(4), 464–471. https://doi.org/10.1159/000374090.Acetaminophen
7) Tuñón, M. J., Alvarez, M., Culebras, J. M., González-gallego, J., Tuñón, M. J., Culebras, J. M., & González-, J. (2009). An overview of animal models for investigating the pathogenesis and therapeutic strategies in acute hepatic failure, *15*, 3086–3098. https://doi.org/10.3748/wjg.15.3086
8) Zakikhan, K., Ph, D., Pournasr, B., Ph, D., Vosough, M., Ph, D., … Ph, D. (2017). In Vitro Generated Hepatocyte-Like Cells : A Novel Tool in Regenerative Medicine and Drug Discovery, (6), 204–217. https://doi.org/10.22074/cellj.2016.4362.Introduction

9) Coronet, P. P., & Road, B. (n.d.). Quantitative Structure Activity Relationship (QSAR), (287).
10) Pagare, A. H., Kankate, R. S., & Shaikh, A. R. (2015). 2D and 3D QSAR using kNN-MFA method of the novel 3 , 4- dihydropyrimidin-2 (1H) -one urea derivatives of N-aryl urea as an antifungal agents ., *5*(2), 1473–1481.
11) Knöspel, F., Jacobs, F., Freyer, N., Damm, G., Bondt, A. De, Wyngaert, I. Van Den, … Zeilinger, K. (n.d.). In Vitro Model for Hepatotoxicity Studies Based on Primary Human Hepatocyte Cultivation in a Perfused 3D Bioreactor System, 1–21. https://doi.org/10.3390/ijms17040584
12) Vinken, M., Jaeschke, H., Brussel, V. U., & City, K. (2017). HHS Public Access, 86–97. https://doi.org/10.1016/j.taap.2015.11.016.Experimental
13) Karelson, M., Karelson, G., Tamm, T., Tulp, I., &Jänes, J. (2009). QSAR study of pharmacological permeabilities, *2009*(ii), 218–238.
14) Jaeschke, H., Gores, G. J., Cederbaum, A. I., Hinson, J. A., Pessayre, D., &Lemasters, J. J. (2002). Mechanisms of hepatotoxicity. *Toxicological Sciences*, *65*(2), 166–176. https://doi.org/10.1093/toxsci/65.2.166
15) Piparo, E. Lo, & Worth, A. (2010). *Review of QSAR Models and Software Tools for predicting Developmental and Reproductive Toxicity*. https://doi.org/10.2788/9628
16) The Drug Discovery and Development Process 7. (n.d.). *Drug Discovery and Development*, 34597.
17) Kapetanovic, I. M. (2011). *Drug Discovery and Development – Present and Future*. https://doi.org/10.5772/1179
18) Andrade, R. J., Lucena, M. I., Martin-vivald, R., Fernandez, M. C., Nogueras, F., Pelaez, G.,
19) … Sahnern, J. (1999). Acute liver injury associated with the use of ebrotidine, a new Hz-receptor. *Journal of Hepatology*, (5), 641–646. https://doi.org/10.1016/S0168-8278(99)80343-2
20) Pandit, A., Sachdeva, T., &Bafna, P. (2012). Drug-induced hepatotoxicity: A review. *Journal of Applied Pharmaceutical Science*, *2*(5), 233–243. https://doi.org/10.7324/JAPS.2012.2541
21) Mehta, N. (2014). Drug-Induced Hepatotoxicity: Overview, Metabolism of Drugs, Clinical and Pathological Manifestations of Drug-Induced Liver Disease. *Medscape*. Retrieved from http://emedicine.medscape.com/article/169814-overview
22) Björnsson, E. S. (2016). Hepatotoxicity by drugs: The most common implicated agents.
23) *International Journal of Molecular Sciences*, *17*(2). https://doi.org/10.3390/ijms17020224
24) Russmann, S., Kullak-Ublick, G. A., &Grattagliano, I. (2009). Current Concepts of Mechanisms in Drug-Induced Hepatotoxicity. *Current Medicinal Chemistry*, *16*, 3041–3053. https://doi.org/10.2174/092986709788803097
25) Zuch de Zafra, C. L., Ashkenazi, A., Darbonne, W. C., Cheu, M., Totpal, K., Ortega, S., … Dybdal, N. O. (2016). Antitherapeutic antibody-mediated hepatotoxicity of recombinant human Apo2L/TRAIL in the cynomolgus monkey. *Cell Death & Disease*, *7*(8), e2338. https://doi.org/10.1038/cddis.2016.241
26) Burdick, A. D., Sciabola, S., Mantena, S. R., Hollingshead, B. D., Stanton, R., Warneke, J. A., … Whiteley, L. O. (2014). Sequence motifs associated with hepatotoxicity of locked nucleic acid—modified antisense oligonucleotides. *Nucleic Acids Research*, *42*(8), 4882–4891. https://doi.org/10.1093/nar/gku142
27) El-Hosseiny, L. S., Alqurashy, N. N., &Sheweita, S. A. (2016). Oxidative stress alleviation by sage essential oil in Co-amoxiclav induced hepatotoxicity in rats. *International Journal of Biomedical Science*, *12*(2), 71–78.
28) Talib, W. H., &Abukhader, M. M. (2013). Combinatorial effects of thymoquinone on the anticancer activity and hepatotoxicity of the prodrug CB 1954. *Scientia Pharmaceutica*, *81*(2), 519–530. https://doi.org/10.3797/scipharm.1211-15
29) Moeller, T. A., Shukla, S. J., & Xia, M. (2012). Assessment of Compound Hepatotoxicity Using Human Plateable Cryopreserved Hepatocytes in a 1536-Well-Plate Format. *ASSAY and Drug Development Technologies*, *10*(1), 78–87. https://doi.org/10.1089/adt.2010.0365
30) Chen, Q., Xue, Y., & Sun, J. (2013). Kupffer cell-mediated hepatic injury induced by silica

nanoparticles in vitro and in vivo. *International Journal of Nanomedicine*, *8*, 1129–1140. https://doi.org/10.2147/IJN.S42242

Supplementary Material:

1 List of withdrawn drugs (As Test Set)

SR no.	Compound_name	Predicted Value **[-Log10(mol/kg)**
1	Busulfan	2.81
2	amiodaron	3.52
3	EXIFONE	2.28
4	FIPEXIDE	2.51
5	Pemoline	2.20
6	Nefazodone	2.66
7	NIALAMIDE	2.49
8	Amineptine	2.95
9	Etifoxine	2.97
10	Tolcapone	2.00
11	NITREFAZOLE	2.42
12	ALPIDEM	2.48
13	Clomacran	2.91
14	Fenclozic acid	2.60
15	Mebanazine	3.12
16	Ibufenac	2.18
17	Phenoxypropazine	2.37
18	Cinchophen	1.71
19	Chlormezanone	2.74
20	Bendazac	2.42
21	Lumiracoxib	3.03
22	Pirprofen	3.13
23	Benoxaprofen	2.88
24	Droxicam	2.77
25	Trovafloxacin	2.48
26	Temafloxacin	2.26
27	Xenazoic acid	3.75
28	Alatrofloxacin	2.55
29	Ketoconazole	3.05
30	Dilevalo	2.82
31	Suloctidil	2.14
32	Moxisylyte	2.40
33	Ticrynafen	2.83
34	Sitaxentan	2.82
35	Beclobrate	2.26
36	Perhexiline	2.45
37	Benzarone	2.32
38	Tolrestat	3.10
39	Troglitazone	2.79
40	Oxyphenisatine	2.69
41	Pipamazine	2.76
42	PHENISATIN	3.02
43	Ebrotidine	2.72

2. List of Approved Drugs (As training set)

SR no.	Compound_name	Predicted Value [-Log10(mol/kg)]
1	Verapamil	3.18
2	Clonidine	2.97
3	Losartan	2.58
4	Tazobactam	2.94
5	Cefuroxime	2.38
6	Amlodipine	2.38
7	Pethidine	2.85
8	Hydralazine	2.79
9	Candesartan cilexetil	2.31
10	Tranexamic Acid	1.85
11	Duloxetine	2.46
12	Lorpiprazole	3.53
13	Viomycin	1.70
14	Danazol	2.36
15	Miglitol	1.11
16	Minaprine	2.67
17	Barnidipine	2.80
18	Ivacaftor	2.20
19	Darunavir	2.57
20	Sotalol	2.47
21	Ergotamine	4.34
22	Metolazone	2.07
23	Cidofovir	1.87
24	Difluprednate	3.42
25	Sulfoxone	2.07
26	Cilazapril	2.48
27	Arbutamine	2.45
28	Anagrelide	3.58
29	Ethambutol	2.11
30	Atomoxetine	2.16
31	Oxyphencyclimine	3.09
32	Cevimeline	2.19
33	Aminophylline	3.02
34	Efinaconazole	2.34
35	Chloropyramine	2.68
36	Indacaterol	2.90
37	Cycrimine	2.69
38	Toloxatone	2.59
39	Elvitegravir	2.71
40	Methyltestosterone	2.15
41	Fludarabine	2.19
42	Idarubicin	2.80
43	Ibrutinib	2.74
44	Crizotinib	2.46
45	Ruxolitinib	2.33
46	Ziprasidone	2.60
47	Metocurine	2.08
48	Stiripentol	2.43

49	Trihexyphenidyl	2.78
50	Pamidronate	1.52
51	Pirlindole	2.70
52	Lorpiprazole	3.53
53	Fenspiride	2.61
54	Metixene	2.69
55	Isocarboxazid	3.03
56	Acyclovir	2.63