ECMC-P 2025 Conference
The 1st International Electronic Conference on Medicinal Chemistry and Pharmaceutics
01–30 November 2025 | Online
MDPI

Machine Learning Based Classification Analysis of hERG Blockers for Prevention of Cardiac Toxicity

Sushmita Verma[1], C. Karthikeyan[1], N.S. Hari Narayana Moorthy[1]*

[1]Computational and Theoretical Medicinal Chemistry Laboratory, Department of Pharmacy, Indira Gandhi National Tribal University, Amarkantak, Madhya Pradesh-484887, India

INTRODUCTION & AIM

The human ether-a-go-go-related gene (hERG) (K+ channel protein) predominantly expresses in cardiac and nervous tissues, and its blockade by small molecule drugs is linked to prolonged QT intervals, resulting in long QT syndrome (LQTS). They induce ventricular tachyarrhythmia known as torsades de pointes (TdP) and sudden cardiac death. This cardiac toxicity is one of the main problems in recent drug development and discovers a drug without blockade of the hERG channel.

Medicinal chemists are working towards for discovering bioactive molecules without hERG blockage. Analysis of the structural features responsible for the hERG blocking activity of drugs can provide information about the features required for this activity. Hence, many strategies have been implemented in last few decades including computational techniques for the design of novel molecules and the machine learning based classification analysis is predominantly used nowadays. The available literature showed that QSAR and computational analyzes were reported in the series of compounds, hence classification analysis has been performed to analyze the structural features required for the hERG blockade.

METHOD

Data Set Processing and Descriptors Calculation:

A data set comprised of 451 hERG inhibitors with diverse structures, obtained from literature and binding database was considered for the present classification analysis. Initially, the inhibitory concentration in molar units was converted into -logIC_{50} or pIC_{50} to reduce the skewness of the data and for correlating the concentration to the free energy changes of the molecules. The pIC_{50} values varied between 0 to 9.6. The value of 5.5 was kept as a threshold value to discriminate between active and inactive compounds, because pIC_{50} values for most of the compounds in the series are distributed between 4 and 7. The chemical structures of the drugs were drawn and its physicochemical descriptors (1444 1D and 2D physicochemical descriptors) and Molecular ACCess Systems Fingerprints (166 MACCSFP) were calculated using the PaDEL software (padel.nus.edu.sg/software).

Machine Learning Studies:

Random Forest (RF), Random Tree (RT) and Decision Tree (DT) algorithms in the Waikato Environment for Knowledge Analysis (WEKA) software were utilized for this classification analysis. The physicochemical properties and MACCS fingerprints were considered as independent variables and the hERG blocking activity of the compounds were treated as dependent variables. The statistical parameters with significant values were considered to select best models using the complete dataset (training). The models were validated using a test set (30% of the data for training) and the 10-fold cross-validation methods.

RESULTS & DISCUSSION

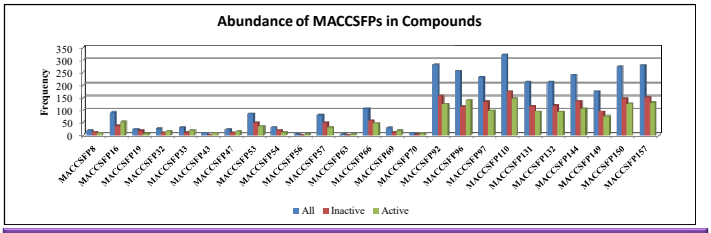
The results obtained from the analysis showed that the models developed with complete training set compounds and validated by test set (30%) and 10-fold cross-validation methods provided significant statistical parameters such as Sensitivity, Specificity, Accuracy, Precision, Enrichment Factor (EF), G-Mean, Mathew's correlation coefficients (MCC), ROC, etc

Classification results of the models

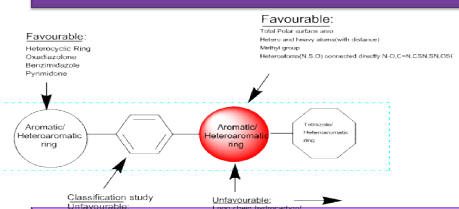
Statistical Parameters	Model A-RF	Model B-RF	Model C-RF	Model A-RT	Model B-RT	Model C-RT	Model A-DT	Model B-DT	Model C-DT
No. Descriptors	30	30	30	10	10	10	14	14	14
Total P (TP/FN)	191	51	191	191	51	191	222	66	222
Total N (TN/FP)	257	83	257	257	83	257	228	69	228
Sensitivity	0.89	0.67	0.61	0.60	0.49	0.51	0.73	0.59	0.53
Specificity	0.97	0.73	0.74	0.87	0.67	0.76	0.92	0.67	0.64
Accuracy	0.94	0.71	0.68	0.75	0.60	0.66	0.83	0.63	0.59
Precision	0.955	0.607	0.630	0.772	0.481	0.616	0.901	0.629	0.590
NPV	0.9222	0.7821	0.7159	0.746	0.683	0.678	0.781	0.630	0.584
FNR	0.110	0.333	0.393	0.398	0.510	0.487	0.266	0.409	0.469
FPR	0.031	0.265	0.265	0.132	0.325	0.237	0.079	0.333	0.360
FDR	0.449	0.393	0.370	0.228	0.519	0.384	0.944	0.371	0.41
Informedness	0.86	0.40	0.34	0.47	0.16	0.27	0.65	0.26	0.18
Markedness	0.87	0.39	0.35	0.52	0.16	0.29	0.68	0.26	0.18
AUC	0.93	0.70	0.67	0.73	0.58	0.64	0.83	0.63	0.59
MCC	0.86	0.39	0.33	0.48	0.16	0.28	0.66	0.25	0.18

The contributed descriptors explained that the hERG blockers contain aromatic rings, such as phenyl (C_6H_5) or benzyl ($C_6H_5CH_2$) groups, which can participate in π-π interactions with aromatic residues in the channel. Basic nitrogen atoms, often found in primary (R-NH_2), secondary (R_2-NH), or tertiary (R_3-N) amines, are common in hERG blockers. These amines can form ionic interactions with acidic residues, like glutamate (Glu) or aspartate (Asp), in the channel. Hydrophobic groups, such as alkyl chains (e.g., CH_3, C_2H_5) or cycloalkyl groups (e.g., cyclohexyl, C_6H_{11}), play a significant role by enhancing binding affinity to the hydrophobic pocket of the hERG channel. The presence of flexible linkers, such as ethylene (–CH_2–CH_2–) or propylene (–CH_2–CH_2–CH_2–) chains, between hydrophobic and hydrophilic parts of a molecule can facilitate optimal orientation within the binding site. Additionally, hydrogen bond donors (e.g., –OH, –NH_2) and acceptors (e.g., –CO, –CN) contribute to the binding affinity and specificity of a compound for the hERG channel, although they are less critical than hydrophobic and ionic interactions.

Frequency of MACCS fingerprints contributed in the model building and activity classification" grouped bar chart.



CONCLUSION



The figure showed the favourable and unfavourable conditions for hERG toxicity of the molecules. The compounds with polar functional groups are highly contributed for the hERG blocking activity.

FUTURE WORK / REFERENCES

The novel bioactive molecules may be developed using these results, which shall be free from cardiac toxicity.

•Moorthy NSHN, Ramos MJ, Fernandes PA. Predictive QSAR models development and validation for human ether-a-go-go related gene (hERG) blockers using newer tools. 2013;6366:1–8.
•Singh S, Karthikeyan C, Moorthy NSHN. Classification analysis of fatty acid synthase inhibitors using multialgorithms on topological descriptors and structural fingerprints. Chemical Biology and Drug Design. 2023;101(2):395–407