Assessment of Drug Proarrhythmicity Using Artificial Neural Network with *in Silico* Deterministic Model Outputs

Yedam Yoo
Department of IT Convergence Engineering, Kumoh Institute of Technology, Gumi
Aroli Marcellinus
Department of IT Convergence Engineering, Kumoh Institute of Technology, Gumi
Da Un Jeong
Department of IT Convergence Engineering, Kumoh Institute of Technology, Gumi
Ki-Suk Kim
R&D Center for Advanced Pharmaceuticals & Evaluation, Korea Institute of Toxicology, Korea Research Institute of Chemical Technology, Daejeon
Ki Moo Lim (kmlim@kumoh.ac.kr)
Department of IT Convergence Engineering, Kumoh Institute of Technology, Gumi

Research Article

Keywords: Drug, Neural Network, CiPA, ICH

Posted Date: March 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-244771/v1

Abstract

Methodologies for predicting the occurrence of torsade de pointes by drugs via computer simulations have been developed and verified recently, as part of the Comprehensive *in vitro* Proarrhythmia Assay initiative. However, the predictive performance still requires improvement. Herein, we propose a deep learning algorithm based on artificial neural networks that receives nine multiple features and considers the action potential morphology, calcium concentration morphology, and charge characteristics to further improve drug toxicity evaluation performance. The voltage clamp experimental data for 28 drugs were augmented to 2,000 data entries using an uncertainty quantification technique. By applying these data to the modified Ohara Rudy *in silico* model, nine features (dV_m/dt_{max}, $AP_{resting}$, APD90, APD50, $Ca_{resting}$, CaD90, CaD50, qNet, and qInward) were predicted. These nine features were used as inputs to an artificial neural network (ANN) model to classify drug toxicity into high-risk, intermediate, and low-risk groups. The model was trained with data of 12 drugs and tested with the data of the remaining 16 drugs. The proposed ANN model demonstrated an AUC of 0.94 in the high-risk group, 0.73 in the intermediate group, and 0.91 in the low-risk group. This is higher than the classification performance of the method proposed in previous studies.

Introduction

In 1999, the gastro prokinetic agent, cisapride, was expelled from the European pharmaceutical market due to torsades de pointes (TdP)[1]. In 2005, the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) established guidelines for the proarrhythmic assessment of the drug[2]. This guideline suggested that the cardiotoxicity assessment for drugs should be conducted according to the S7B nonclinical evaluation and the E14 clinical evaluation guidelines. This conventional guideline requires extensive trials and has high sensitivity but low specificity for the risk classification of drugs. This means that even drugs that do not undergo TdP are strictly regulated, negatively affecting drug development[3].

The Comprehensive *in vitro* Proarrhythmia Assay (CiPA) project was established with 13 advanced medical institutions attending the think-tank conference hosted by the FDA headquarters in 2013 to revise the existing drug development guidelines. The main change in the S7B nonclinical evaluation guideline through the CiPA project is the evaluation of the drug response of multiple ion channels using the *in silico* method from the hERG channel single analysis evaluation method through *in vitro* experiments[4, 5].

Dutta et al. proposed an *in silico* model that modified the ORD model. This optimized the maximum conductivities of the IKs, ICaL, IKr, INaL, and IK1 ion channels to 1.870, 1.007, 1.013, 2.661, and 1.698, respectively. This model corrected the maximum conductance of underestimated or overestimated channels, allowing the drug response in the *in silico* model to be simulated similarly to that *in vitro*. By deriving the qNet (sum of the area under the current graph over time of 6 ion channels–INaL, ICaL, IKr, Ito, IK1, and IKs) using the *in silico model*, the criteria for classifying the risk of TdP occurrence of drugs into high-, intermediate-, and low-risk levels were established[6].

Parikh et al. performed simulations by applying drug effects to the Ord model to evaluate the drug response of multiple ion channels in a complex manner. They extracted a group of TdP-inducing drugs from the derived results using a logical regression technique. The 13 electrophysiological characteristics (upstroke velocity, peak voltage, APD50, APD @−60 mV, APD90, Resting voltage, AP triangulation, diastolic [Ca2+]i, the amplitude of CaT, peak [Ca2+]i, CaTD50, CaTD90, CaT triangulation) used as inputs to the logistic regression model are derived from *in silico* simulation under the condition of effective free therapeutic plasma concentration (EFTPC) and drug concentration when blocking the IKr channel by 60%, respectively. Under the drug concentration condition that blocks the IKr channel by 60%, the predictive performance of early after depolarization (EAD) occurrence exhibited a minimum of 76 points and a maximum of 100 points[7].

Li et al. classified the risk of drugs using qNet as an input to a logical regression model. qNet was calculated using the model that added the hERG dynamic model to the modified ORD model.[6].

Risk groups of drugs were classified by two qNet thresholds obtained through two binary classifications. The accuracy of classification was improved by incorporating the hERG-dynamic model. However, the disadvantages included a large amount of data processing, and mathematical complexity, such as *in vitro* experiments for parameter evaluation and quantification of uncertainty.

Therefore, Liopis-Lorente et al. proposed a new method to classify drug risk groups using nine decision trees. Three features are used as input. The first is T_x (Romero et al., 2018), which is the ratio of the drug concentration when the APD90 increases by 10% and the EFTPC concentration of each drug calculated using the model proposed by Dutta et al. The second feature, T_{qNet}, is the ratio of the calculated qNet value at 10 times the EFTPC concentration and the qNet value at steady state. The third feature, T_{triang}, is the ratio of the difference between APD90 and APD30 calculated at 10 times the concentration of EFTPC and calculated at steady state. The classification accuracy of drug risk groups was 0.899 when T_x was used as an input, 0.908 when T_{triang} was used, and 0.917 in the case of T_{qnet}[8].

In previous studies, researchers typically derived the action potential morphology characteristics, such as APD, or charge characteristics, such as qNet, which can be said to be highly correlated with TdP, from the ORD *in silico* model. The traditional binary classification method was used to classify linear patterns, such as logistic regression and decision trees, using the derived characteristics.[9] However, there is no guarantee that the toxic effects of drugs should be linearly distributed according to risk groups in the action potential and charge characteristics. In addition to the action potential morphology and charge characteristics, the morphology of the calcium concentration also displays a high correlation with the TdP[10].

Therefore, in this study, in order to further increase the accuracy of drug toxicity assessment, we propose a deep learning algorithm based on an artificial neural network. This network has nine multiple characteristic values that consider all of the action potential morphology, calcium concentration morphology, and charge characteristics.

Results

The representative values of AUCs obtained through 10,000 tests using the learned ANN classifier developed in this study were 0.958 for the high-risk level, 0.73 for the intermediate level, and 0.927 for the low risk level. (The median value in the histogram in Fig. 2 is the representative value of AUC, and the 95% range of the confidence interval of the dataset is the verification range). As the classification accuracy of the logistic regression model presented by Li et al., the representative AUC of the high-risk level was 0.856 and the representative AUC of the low-risk level was 0.86 (Methodologically, the AUC of the intermediate level could not be predicted). Therefore, the accuracy of the ANN classifier developed in this study was 10.2% higher at the high-risk level and 6.7% higher at the low risk level than the accuracy suggested in a previous study (Li et al. 2018). The minimum value of the confidence interval for the ANN classifier was 9% higher at the high-risk level and 7% higher at the low risk level, and the maximum value in the confidence interval for the ANN classifier was 10% higher at the high-risk level and 10.5% at the low risk level (Table 2).

Discussion

In this study, a deep learning algorithm based on artificial neural networks was developed to evaluate drug toxicity by inputting nine multiple characteristic values, including action potential morphology (APD90, APD50, dV/dt_{max}, and $AP_{resting}$), calcium concentration morphology (CaD90, CaD50, and $Ca_{resting}$), and charge characteristics (qNet and qInward). An *in silico* simulation using the Dutta model was performed to derive nine characteristic values. As for the performance of the classification algorithm, when comparing the results through the same *in silico* model and validation method as performed by Li et al., our performance was 10.2% superior at the high risk level and 6.7% superior in the low risk level than the classification performance presented by Li et al. In addition, by using an artificial neural network instead of the logistic regression classification method used by Li et al., it was possible to explicitly classify not only high-and low-risk levels, but also intermediate levels.

A limitation of this study is that nine parameters should be provided as input values for the toxicity assessment classifier. This is significantly more than the one input value required in previous studies. To obtain nine reliable parameters, sufficient reliability of the physiological/pharmacological *in silico* model of cardiomyocytes should be supported. A second limitation is that in previous studies, risk groups were classified through threshold values of physiologically/pharmacologically meaningful parameters, such as qNet, qInward, and TqNet. However, the artificial neural network model proposed in this study does not provide an explicit threshold for classification. This means that when a researcher developing a new drug uses this artificial neural network classifier for cardiac toxicity assessment, it is difficult to causally evaluate the validity of the classified results. The above two limitations are inevitably encountered if an artificial neural network-based deep learning method is used. Nevertheless, it is highly meaningful in the field of new drug development research to develop an algorithm with higher toxicity assessment and classification performance than in previous studies.

Methods

Software and Data: Hill fitting and Bootstrap

For this study, the same data fitting method and *in silico* model used by Li et al. were re-implemented based on the C++ language. We used the patch-clamp experiment data uploaded to the GitHub website (https://github.com/FDA/CiPA/) from the CiPA project group. Hill fitting was performed using the experimental data obtained by Crumble et al.[5] through patch clamp to determine the effect of the drug on six ion channels. To quantify the uncertainty of the experimental data, we extracted 2,000 Hill coefficients and IC50s for six ion channels by bootstrapping within 95% of the confidence interval using the MCMC model proposed by Chang et al. (2017). [11]. Hill coefficients and IC50 modify the conductivity of six ion channels, and the conductivity of these modified ion channels is applied to the *in silico* simulation(Fig.1).

In silico simulation protocol

The Ohara Rudy model modified by Dutta et al. was used as an electrophysiological *in silico* model of cardiomyocytes. We used 28 drugs in a total of 8 high-risk groups, 11 intermediate groups, and 9 low-risk groups according to the risk of cardiac arrhythmia induction (Table 1), and applied the 2,000 IC50s and Hill coefficients obtained from each drug through Hill fitting and bootstrap to the *in silico* model[12]. The concentration of the drug was set to the maximum serum concentration (free Cmax), which is the characteristic value of each drug, and the concentrations corresponding to 2, 3, and 4 times that of each drug. In total, the simulation was 1,000 stimulations applied with a stimulation period of 2 s, stimulation intensity of -80 μA/μF, and stimulation duration of 0.1 ms for 2,000 possible drug-affected *in silico* models under four different concentrations of each drug.

Feature Evaluation

Nine features related to TdP are derived through a single-cell electrophysiology simulation. Features include action potential characteristics, calcium characteristics, and ion charge characteristics. Among the action potential characteristics, the action potential duration 90 (APD90) is the duration between the depolarization point and the repolarization point 90% below the maximum amplitude in the shape of the action potential. APD50 is the duration between the depolarization and repolarization points 50% below the maximum amplitude in the action potential shape. dV_m/dt_{max} is the maximum slope when the membrane potential is depolarized in the shape of the action potential, and $AP_{resting}$ is the resting membrane potential. Calcium characteristics include calcium transient duration 90 (CaD90), which is the duration between 90% or less of the maximum amplitude during the transient period of intracellular calcium. CaD50 is the duration between 50% or less of the maximum amplitude during the intracellular calcium transient. $Ca_{resting}$ is the diastolic concentration of intracellular calcium. The qNet of the ion charge characteristic is the total amount of ion charges that passed by the six ion channels (INaL, ICaL, IKr, IKs, IK1, Ito) until the end of AP, and is calculated as the sum of the integral of the current graph over time (equation 1). qNet was described by Li et al. (2018) and was used as an input feature for classifying

the risk of TdP-induced drugs using a logistic regression model performed by Li et al. qInward is the amount of charge changed through the ICaL and INaL ion channels during the AP beat induced by the drug[11]. (equation 2):

$$\boldsymbol{qnet} = \int_0^{BCL} (INaL + ICaL + IKr + IKs + Ik1 + Ito)dt \quad \textbf{equation (1)}$$

$$\boldsymbol{qinward} = (ICaL\_drug\_AUC/ICaL\_control\_AUC + INaL\_drug\_AUC/INaL\_control\_AUC)/2 \quad \textbf{equation (2)}$$

_drug_AUC is the area under the current change graph over time of each ion channel when the drug condition and _control_AUC is the area under the current change graph over time of each ion channel without drug conditions. The action potential selection criterion for feature calculation is one action potential with the highest dVm/dtrepol value in repolarization, except that depolarization or repolarization fails for the last 250 action potentials from 1,000 action potentials. Nine features were calculated from one selected action potential. An *in silico* simulation was performed for each drug concentration (Cmax, Cmax*2, Cmax*3, Cmax*4), and the average value of the calculated nine features was assigned as the input of the ANN model.

Artificial Neural Network Model

The proposed ANN model is composed of an input layer with nine nodes that consider nine features (dvm/dtmax, APresting, APD90, APD50, Caresting, CaD90, CaD50, qNet, and qInward) as inputs; a hidden layer with five nodes; and an output layer with three nodes(Fig.1). “MinMaxScaler” was used for input data to avoid overfitting or underfitting caused by unit differences between features. To prevent overfitting during the training process, 10-fold cross-validation was performed. “ReLu” was used as the activation function of the hidden layer, and “Softmax” was used as the activation function of the output layer. The output layer node contained multiple classifications: high, intermediate, and low risk levels.

The data for ANN learning were acquired through *in silico* simulation. Since 2,000 samples were obtained for each drug and 12 drugs were obtained, a total of 24,000 samples were used. Data for the ANN test was obtained from 16 drugs. The dataset used for testing the model contains 16 drug data, and is randomly selected from 2,000 data points. The test is performed 10,000 times using 10,000 data sets containing data from 16 test drugs individually. As a test result, the performance of the model was evaluated using the area under the curve (AUC) corresponding to the area of the receiver operating characteristic (ROC curve). Here, 95% of the AUC confidence interval and the median of the frequency distribution of 10,000 AUC results were set as representative values for comparison of the results.

Declarations

Acknowledgements (not compulsory)

This research was partially supported by NRF (National Research Foundation) under basic engineering research project (2016R1D1A1B0101440) and the EDISON (NRF-2011-0020576) Programs, and the

Grand Information Technology Research Center Program through the Institute of Information & Communications Technology and Planning & Evaluation (IITP) funded by the Ministry of Science and ICT (MSIT), Korea (IITP-2021-2020-0-01612).

Author contributions statement

This manuscript is the intellectual product of the entire team. YDY wrote the machine learning source code and the manuscript, performed the data analysis, and interpreted the results. KML designed the study, created simulation code, and reviewed and revised the entire manuscript based on the simulation results. All authors have read and approved the final manuscript.

Data availability

All datasets used in the study were generated through simulations performed by the authors based on the methods described in the text.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Roden, D. M. Cellular basis of drug-induced torsades de pointes. *Br. J. Pharmacol.***154**, 1502–1507 (2008).
2. Cavero, I. & Crumb, W. ICH S7B draft guideline on the non-clinical strategy for testing delayed cardiac repolarisation risk of drugs: A critical analysis. *Expert Opin. Drug Saf.***4**, 509–530 (2005).
3. Colatsky, T. *et al.* The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress. *J. Pharmacol. Toxicol. Methods.***81**, 15–20 (2016).
4. Oh, K., Kim, K., Lee, H., Han, S. & Yim, D. Integrated in vivo cardiac safety evaluation using systemic pharmacology technique. **12**,25–32(2018).
5. Crumb, W. J., Vicente, J., Johannesen, L. & Strauss, D. G. An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel. *J. Pharmacol. Toxicol. Methods.***81**, 251–262 (2016).
6. Dutta, S. *et al.* Optimization of an in silico cardiac cell model for proarrhythmia risk assessment. *Front. Physiol.***8**, 1–15 (2017).
7. Fermini, B. *et al.* A new perspective in the field of cardiac safety testing through the comprehensive in vitro proarrhythmia assay paradigm. *J. Biomol. Screen.***21**, 1–11 (2016).
8. Llopis-Lorente, J. *et al.* InSilico Classifiers for the Assessment of Drug Proarrhythmicity. *J. Chem. Inf. Model.***60**, 5172–5187 (2020).

9. Ismail, W. N., Hassan, M. M., Alsalamah, H. A. & Fortino, G. CNN-based health model for regular health factors analysis in internet-of-medical things environment. *IEEE Access.***8**, 52541–52549 (2020).
10. Golgooni, Z. *et al.* Deep learning-based proarrhythmia analysis using field potentials recorded from human pluripotent stem cells derived cardiomyocytes.*IEEE J. Transl. Eng. Heal. Med.***7**, (2019).
11. Chang, K. C. *et al.* Uncertainty quantification reveals the importance of data variability and experimental design considerations for in silico proarrhythmia risk assessment. *Front. Physiol.***8**, 1–17 (2017).
12. Li, Z. *et al.* Assessment of an In Silico Mechanistic Model for Proarrhythmia Risk Prediction Under the CiPA Initiative. *Clin. Pharmacol. Ther.***105**, 466–475 (2019).
13. Li, Z. *et al.* Improving the in silico assessment of proarrhythmia risk by combining hERG (Human Ether-à-go-go-Related Gene) channel-drug binding kinetics and multichannel pharmacology. *Circ. Arrhythmia Electrophysiol.***10**, 1–12 (2017).

Tables

Table 1. Twenty-eight drugs selected by the CiPA research group classified into high, intermediate, and low risk levels according to the possibility of causing TdP[3]. Twelve drugs were used in machine learning training and sixteen drugs were used in testing**.**

Used \ risk level	High	Intermediate	Low
Training	Quinidine	Cisapride	Verapamil
	Sotalol	Terfenadine	Ranolazine
	Dofetilide	Chlorpromazine	Diltiazem
	Bepridil	Ondansetrom	Mexiletine
Testing	Disopyramide	Clarithromycin	Metoprolol
	Ibutilide	Clozapine	Nifedipine
	Vandetanib	Domperidone	Nitrendipine
	Azimilide	Droperidol	Tamoxifen
		Pimozide	Loratadine
		Risperidone	
		Astemizole	

Table 2 Comparison of the accuracy of prediction of TdP-induced risk levels when using an artificial neural network (ANN) model and logical regression model proposed by Li et al.

Model	High	Intermediate	Low
Logistic regression [12]	0.86(0.81-0.9)	-	0.86(0.82-0.90)
ANN	0.94(0.75-1.00)	0.73(0.52-0.91)	0.91(0.73-1.00)

Figures

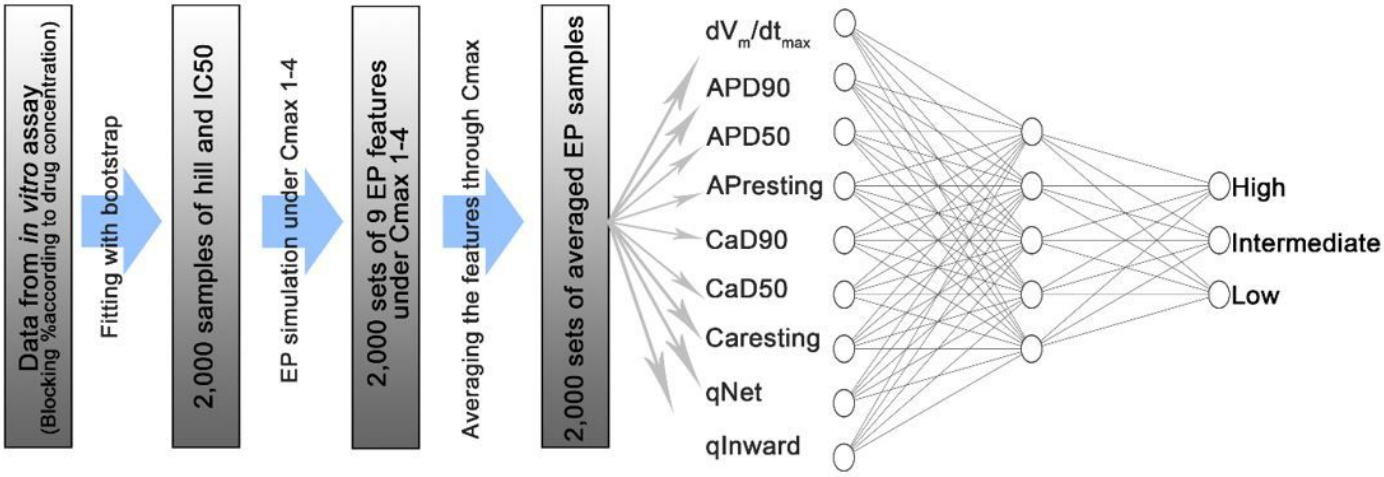

Figure 1

Schematic diagram of an artificial neural network model consisting of an input layer with 9 nodes, a hidden layer with 5 nodes, and an output layer with 3 nodes. Here, dVm/dtmax is the maximum slope when the membrane potential is depolarized in the shape of the action potential; APD90 is the duration between the depolarization point and the repolarization point 90% below the maximum amplitude in the shape of the action potential; APD50 is the duration between the depolarization and repolarization points 50% below the maximum amplitude in action potential shape; APresting is the resting membrane potential; CaD90 is the duration between 90% or less of the maximum amplitude during the transient period of intracellular calcium; CaD50 is the duration between 50% or less of the maximum amplitude during the intracellular calcium transient; Caresting is the diastolic concentration of intracellular calcium; qNet is the total amount of ion charges that move through the six ion channels (INaL, ICaL, IKr, IKs, IK1, Ito) during the action potential duration[6]; qInward is the average of the ratio between the drug reaction and the steady state of charges directed to the cell through the ICaL and INaL ion channels during the action potential period[13].

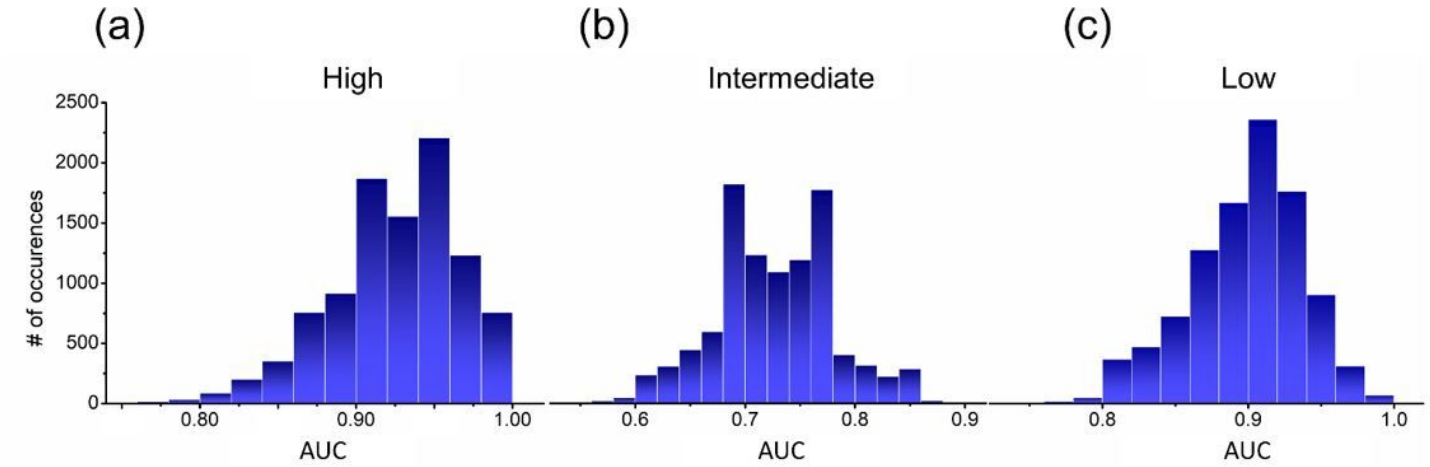

Figure 2

Histogram representing the frequency of AUCs obtained through 10,000 tests. (a) high risk group; (b) intermediate risk group; (c) low risk group