

Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity

Chuipu Cai,^{†,‡,ID} Pengfei Guo,[†] Yadi Zhou,[§] Jingwei Zhou,[†] Qi Wang,[†] Fengxue Zhang,[‡] Jiansong Fang,^{*,†,ID} and Feixiong Cheng,^{*,||,‡,#,ID}

[†]Institute of Clinical Pharmacology, Guangzhou University of Chinese Medicine, Guangzhou 510405, China

[‡]School of Basic Medical Sciences, Guangzhou University of Chinese Medicine, Guangzhou 510405, China

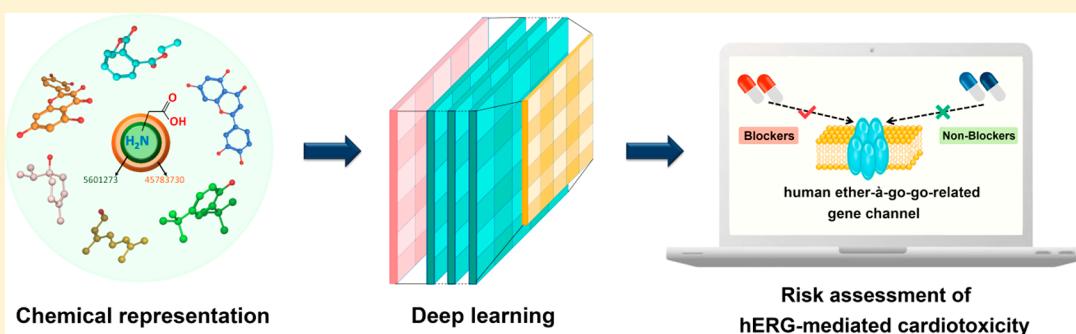
[§]Department of Chemistry and Biochemistry, Ohio University, Athens, Ohio 45701, United States

[¶]Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44106, United States

[⊥]Department of Molecular Medicine, Cleveland Clinic Lerner Research Institute, Case Western Reserve University, 9500 Euclid Avenue, Cleveland, Ohio 44195, United States

[#]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, United States

Supporting Information



ABSTRACT: Blockade of the human ether-à-go-go-related gene (hERG) channel by small molecules induces the prolongation of the QT interval which leads to fatal cardiotoxicity and accounts for the withdrawal or severe restrictions on the use of many approved drugs. In this study, we develop a deep learning approach, termed deephERG, for prediction of hERG blockers of small molecules in drug discovery and postmarketing surveillance. In total, we assemble 7,889 compounds with well-defined experimental data on the hERG and with diverse chemical structures. We find that deephERG models built by a multitask deep neural network (DNN) algorithm outperform those built by single-task DNN, naïve Bayes (NB), support vector machine (SVM), random forest (RF), and graph convolutional neural network (GCNN). Specifically, the area under the receiver operating characteristic curve (AUC) value for the best model of deephERG is 0.967 on the validation set. Furthermore, based on 1,824 U.S. Food and Drug Administration (FDA) approved drugs, 29.6% drugs are computationally identified to have potential hERG inhibitory activities by deephERG, highlighting the importance of hERG risk assessment in early drug discovery. Finally, we showcase several novel predicted hERG blockers on approved antineoplastic agents, which are validated by clinical case reports, experimental evidence, and the literature. In summary, this study presents a powerful deep learning-based tool for risk assessment of hERG-mediated cardiotoxicities in drug discovery and postmarketing surveillance.

INTRODUCTION

The human ether-à-go-go-related gene (hERG) encodes the pore-forming α -subunit of rapid delayed rectifier current, playing crucial roles in the regulation of exchanges of the resting potential and action potential on the cardiac myocyte.^{1,2} Overwhelming experimental and clinical evidence have indicated that a blockade of the hERG channel can induce long-QT syndrome (LQTS), which may lead to fatal cardiotoxicities, such as torsade de pointes (TdP) arrhythmia.³ To date, several drugs, including astemizole, terfenadine, vardenafil, cisapride, and ziprasidone, have been withdrawn or severely restricted on the use for the undesirable hERG-related cardiac side effects.^{4–6} Since hERG channel is highly sensitive to be inhibited by a large amount of structurally diverse

molecules, an early evaluation of hERG blockade has become a necessary step in drug discovery.^{7,8}

According to the guideline published by International Conference of Harmonization, all new drugs should be assessed preclinically for their hERG inhibitory activities before submitted to regulatory reviews.⁹ However, current *in vivo* and *in vitro* methods for screening hERG blockers, such as rubidium-flux assays, fluorescence-based assays, electrophysiology measurements, and radioligand binding assays, are costly,

Special Issue: Machine Learning in Drug Discovery

Received: October 31, 2018

Published: February 4, 2019



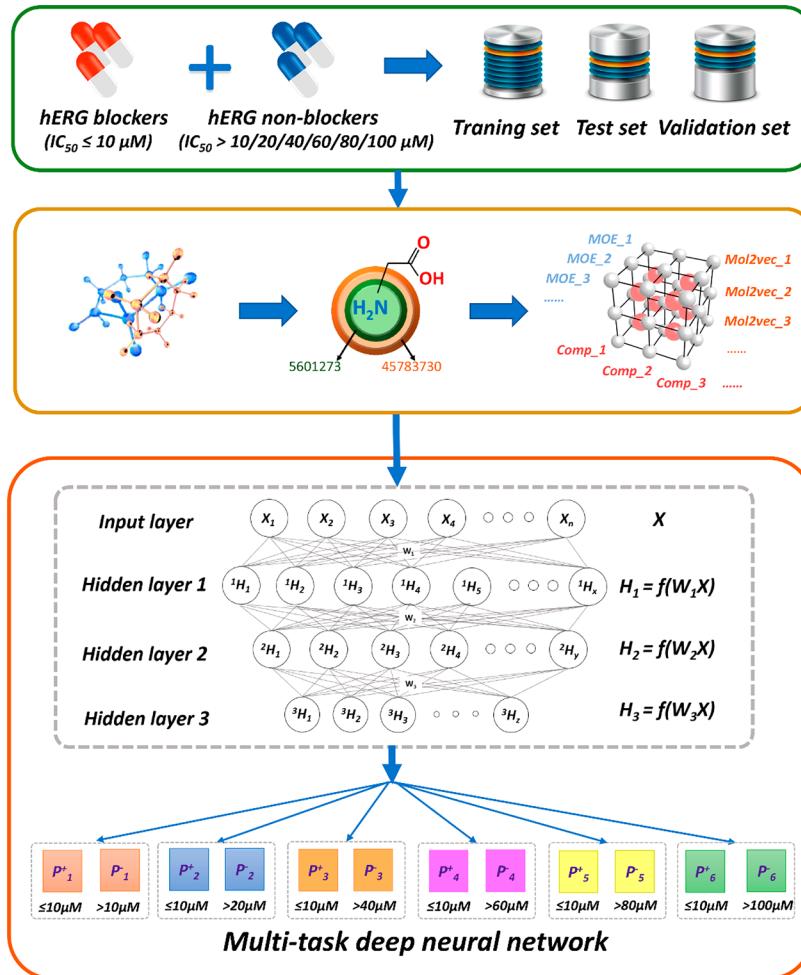


Figure 1. Diagram illustrating the deephERG framework. A comprehensive database on the hERG blockers ($IC_{50} \leq 10 \mu M$) and nonblockers with different decoy thresholds values ($IC_{50} > 10, 20, 40, 60, 80$, and $100 \mu M$, respectively) are divided into training sets, test sets, and validation sets. The deep learning models are built based on Mol2vec and MOE descriptors using a multitask deep neural network containing three hidden layers (see Materials and Methods).

laborious, and time-consuming.¹⁰ Recent advances of *in silico* approaches and tools have offered possibilities for effective evaluation of drug pharmacokinetics and pharmacodynamics (PK/PD) properties at the early stage of drug discovery.^{11–19} Over the past several years, a wide range of prediction models for hERG blockers have been published using various machine learning methods.^{4,6,20–30} For instance, in 2010, Doddareddy and co-workers developed classification models from 2,644 compounds using linear discriminant analysis and support vector machine (SVM) methods to estimate the hERG-related cardiotoxicity.²⁸ The area under the receiver operating characteristic curve (AUC) values of models ranged from 0.89 to 0.94 in 5-fold cross validation.²⁸ In 2016, Wang and co-workers utilized pharmacophore modeling combined with machine learning to build classification models for prediction of hERG active compounds. Accuracy for the hERG active and inactive compounds in the test set reached 83.6% and 78.2%, respectively.²⁹ Although some of these models showed acceptable performance on the training set and test set, a small space of chemical diversities has resulted in a limited applicability domain.²⁸ Meanwhile, most of the studies prepared decoy sets by randomly extracting compounds from the entire chemical database. The unknown experimental

evidence of negative samples may cause a potential false positive rate.

In this study, we proposed a multitask deep neural network (DNN) framework for comprehensive assessment of hERG blockers, termed deephERG (Figure 1). First, we collected comprehensive data on the hERG blockade activities of 7,889 diverse chemicals with well-defined experimental end points. All the chemicals were depicted through integration of Molecular Operating Environment (MOE) descriptors³¹ and Mol2vec descriptors.³² We split the data set according to different decoy threshold values (half maximal inhibitory concentration [IC_{50}] = 10, 20, 40, 60, 80, and $100 \mu M$) for building multitask DNN models. For each task, training set, test set, and validation set were used for model training, optimization, and evaluation of generalization abilities, respectively. After systematic comparison, the multitask DNN models offer the best performance comparing to single-task DNN models, traditional machine learning models, and graph convolutional neural network (GCNN). Finally, we applied the best deephERG model to 1,824 FDA approved small molecule drugs for risk assessment of hERG-related cardiotoxicities.³³ In summary, deephERG offers a powerful tool for cardiotoxic risk assessment in drug discovery and postmarketing surveillance.

Table 1. Detailed Description of All Data Sets Used in This Study

threshold value of decoy (μM)	training set			test set			validation set		
	positive	negative	total	positive	negative	total	positive	negative	total
10	3,485	2,826	6,311	435	354	789	435	354	789
20	3,485	1,755	5,240	435	219	654	435	219	654
40	3,485	863	4,348	435	108	543	435	108	543
60	3,485	644	4,129	435	81	516	435	81	516
80	3,485	469	3,954	435	58	493	435	58	493
100	3,485	380	3,865	435	48	483	435	47	482

MATERIALS AND METHODS

Data Preparation. The original compounds with experimental hERG blockage bioactivities were assembled from various well-defined experimental assays: (i) patch-clamp measurements from ChEMBL bioactivity database; (ii) radioligand binding measurements on mammalian and non-mammalian cell lines; (iii) hERG K⁺ channel binding affinity, and (iv) literature-derived data (Supporting Information, Table S1).^{25,28,29,34} We then implemented three criteria: (i) compounds without well-defined experimental hERG blocking bioactivities were eliminated; (ii) incompatible measuring units were converted to unified IC₅₀ value (μM); (iii) only compounds with IC₅₀ value $\leq 10 \mu\text{M}$ were considered as hERG blockers, while the rest were regarded as “decoy pool” for further screening according to different threshold settings.

All compounds were converted into SMILES format and duplicate ones were removed via comparing their InChI keys. In this study, for any two duplicated compounds with inconsistent inhibitory activity values derived from different assays, the one with a higher IC₅₀ value was used for defining hERG blockers. In addition, compounds were further processed by molecular washing and energy minimizing using MOE 2010 software³¹ for protonating strong bases, deprotonating strong acids, removing inorganic counterions, adding hydrogen atoms, generating stereoisomers, and validating single 3D conformers. Finally, a comprehensive collection consisted of 7,889 compounds with well-defined experimental hERG blocking bioactivities was obtained, and 4,355 of the compounds whose experimental values were less than or equal to 10 μM were regarded as hERG blockers (Supporting Information, Table S2 and Table S3).

Subsequently, all compounds were split into three sets—training set, test set, and validation set—with ratios of 8:1:1 via chemical diversity analysis performed by Tanimoto Coefficient measure based on MACCS fingerprint in MOE 2010. Such a split would assign the chemical structures uniformly and avoid potential data bias. Following standard practice, the training sets were used to build deep learning models, and the validation sets for final evaluation of trained models. It is well-known that the performance of deep neural networks is highly sensitive to the selected hyperparameters;³⁵ thus, test sets were applied for tuning model hyperparameters. In this study, hERG blockers were set as “positive” samples, whereas decoys were defined as “negative” samples. Herein, for more rigorous concern, we evaluated several thresholds of decoys, including 10, 20, 40, 60, 80, and 100 μM for building the negative data sets. The chemical information on the data sets is provided in Supporting Information, Table S4, and the detailed statistics of all data sets used in this study are shown in Table 1.

Chemical Representation. We calculated two different types of descriptors for each compound, including MOE 2010 and Mol2vec descriptors.

Mol2vec Descriptors. Mol2vec³² is an unsupervised machine learning approach to learn vector representations of molecular substructures. It is inspired by the natural language processing technique Word2vec.³⁶ Mol2vec learns vector representations of molecular substructures that are pointing to similar directions for chemically related substructures. Compounds can be encoded as vectors via summing up the vectors of the individual substructures and then fed into modeling approaches for compound properties prediction. In this study, the Mol2vec model was pretrained based on a corpus containing 19.9 million compounds and then utilized to feature new samples. The skip-gram method was used with a window size of 10. Finally, 100-dimensional embeddings were generated for all compounds.

MOE Descriptors. Although Mol2vec can well represent the chemical structures of molecules, it still has some limitations. Molecular properties, which are also of great importance for hERG blocker prediction, cannot be comprehensively reflected by Mol2vec. Thus, as a common approach to depict molecular properties, 185 two-dimensional (2D) molecular descriptors were generated by MOE 2010 software for each compound and concatenated to current Mol2vec descriptors to further enhance the representation. Specifically, 185 2D MOE descriptors cover physical property descriptors, subdivided surface area descriptors, atom count and bond count descriptors, adjacency and distance matrix descriptors, Kier and Hall connectivity and Kappa shape indices descriptors, pharmacophore feature descriptors, and partial charge descriptors. The detailed description of these descriptors can be found in a recent study.³¹

Methods for Model Building. In this study, we evaluated three deep learning network algorithms: multitask deep neural network (DNN), single-task DNN, and graph convolutional neural network (GCNN) available from DeepChem.³⁷ In addition, we also evaluated three traditional machine learning methods: naïve Bayes (NB), support vector machine (SVM), and random forest (RF) implemented by Orange canvas (v3.13.0).³⁸

Multitask Deep Neural Network. A typical multitask DNN is composed of interconnected neurons which are arranged hierarchically as layers. The number of neurons in each layer is referred to as “size”. The input layer comprises the neurons for the input vectors based on molecular structure and properties. In the intermediate layers (also known as hidden layers), each hidden neuron applies a weighted sum of the output of the neurons from the previous layer. The output of a hidden layer can be seen as an abstraction of the features of its previous layer. The last layer outputs the prediction results of the model.³⁹ The processed outputs of each hidden layer are

shared across all learning tasks and then input to separate models for each different task.³⁷ Since multitask networks are trained on the joint data, they try to generalize from the data for multiple tasks. The parameters of the shared layers are encouraged to produce joint representations which share information between the tasks.³⁷

One important advantage of a multitask deep network is that it can handle multiple tasks and simultaneously improves the generalization by leveraging the domain-specific information contained in the training samples of related tasks. Thus, we used different thresholds for decoys, including 10, 20, 40, 60, 80, and 100 μM . Each threshold led to a set of decoys (as negative samples), which was considered as one task when combined with positive samples. The six tasks generated from different thresholds were trained simultaneously by the multitask DNN.

In this study, all multitask DNN models comprised three hidden layers with varying input layer sizes according to the number of input features. The first hidden layer had a size of twice as much as that of the input layer. The second and third layers had halved size as the previous ones. The size of the hidden layers were MOE [400, 200, 100], Mol2vec [200, 100, 50], and MOE+Mol2vec [600, 300, 150]. The Adam⁴⁰ algorithm was used for optimization and L^2 normalization term was used for regularization to avoid overfitting. We selected rectified linear unit (ReLU) as the activation function, which was defined as the positive part of its argument:

$$f(x) = x^+ = \max(0, x) \quad (1)$$

where x is the weighted sum of a neuron.⁴¹

Cross entropy was applied as the loss function for the classification task:

$$E = - \sum_{j=1}^T y_i \ln p_j \quad (2)$$

where y_i is the ground truth for total error of the batch as described in eq 3:

$$E = - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^T y_{ij} \ln p_{ij} \quad (3)$$

The number of epochs is an important hyperparameter for the training. A larger number of epochs may bring overfitting problems, while not enough epochs may lead to underfitting. According to the average loss of multitask DNN models based on experimenting with different batch sizes as shown in Supporting Information, Figure S1, the number of epochs was set to 20, since the average loss converged with relatively smaller average loss and lower oscillation. Hyperparameters, including learning rate, weight decay for the L_2 normalization, dropout rate, and weight initialization, were tuned on the test sets with a combination of random hyperparameter search and manual hyperparameter tuning by the hyperparam_search method.^{37,42}

Single-Task Deep Neural Network (Single-task DNN). An independent single-task neural network for each learning task was trained.⁴³ The only difference between the multitask DNN versus single-task DNN is the number of outputs. In the case that a data set contains only a single task, multitask networks are just single-task network.³⁷ In this study, all parameter settings and architecture of single-task DNN were consistent with those using in multitask DNN.

Graph Convolutional Neural Network (GCNN). GCNN exploits the inherent structure of data that is organized into a graph.³⁵ In this study, GCNN used connectivity graphs of molecules converted by ConvMolFeaturizer of DeepChem as input.³⁷ Then several graph convolutions each consisting of a graph convolution layer, a batch normalization layer, and a graph pool layer were sequentially added and were followed by a fully connected layer. The summation of the feature vectors for all nodes generated a graph feature vector and was fed to a classification output layer.³⁷ We set the batch size and number of epochs consistent with those of multitask network.

In addition, three traditional machine learning algorithms, SVM, NB, and RF were also utilized to construct models using the same data sets for comparison. SVM defines a decision boundary that is expressed as a separating hyperplane on the basis of a linear combination of functions parametrized by support vectors.⁴⁴ NB algorithm is a robust classification approach derived from the Bayes theorem with the strong independence assumption that each attribute contributes equally and independently.⁴⁵ RF is an ensemble algorithm for classification, which creates a large number of decision trees by bootstrapping training samples and gives the predictions by integrating the outputs of the individual trees.⁴⁶ For RF, the number of decision trees in forest is set to 100, the depth of individual trees was set to 3, and the number of attributes considered at each split is equal to the square root of the number of attributes in the input data. For SVM, a “grid-search” approach provided by LibSVM 3.23 package⁴⁷ was applied for seeking the penalty parameter C and kernel parameter γ using 5-fold cross validation. For NB, the relative frequency was used for estimating prior class probabilities from the data.

Model Evaluation. Model performance was assessed in terms of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In addition, five metrics, including sensitivity (SE), specificity (SP), prediction accuracy of blockers (Q+), prediction accuracy of nonblockers (Q-), and overall predictive accuracy (Q) were calculated using the following equations.

$$\text{SE} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{SP} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

$$\text{Q}+ = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Q}- = \frac{\text{TN}}{\text{TN} + \text{FN}} \quad (7)$$

$$\text{Q} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \quad (8)$$

Moreover, the area under the receiver operating characteristic (ROC) curve (AUC), which indicates the ability of a classifier to distinguish between two classes, was also computed.⁴⁸ A perfect model has an AUC value of 1, whereas random classifier has a value of 0.5.

RESULTS

Chemical Diversity Analysis. Prior to comparing the performance of different approaches and models, it is pertinent to verify the diversity of the chemical space of the data sets. A

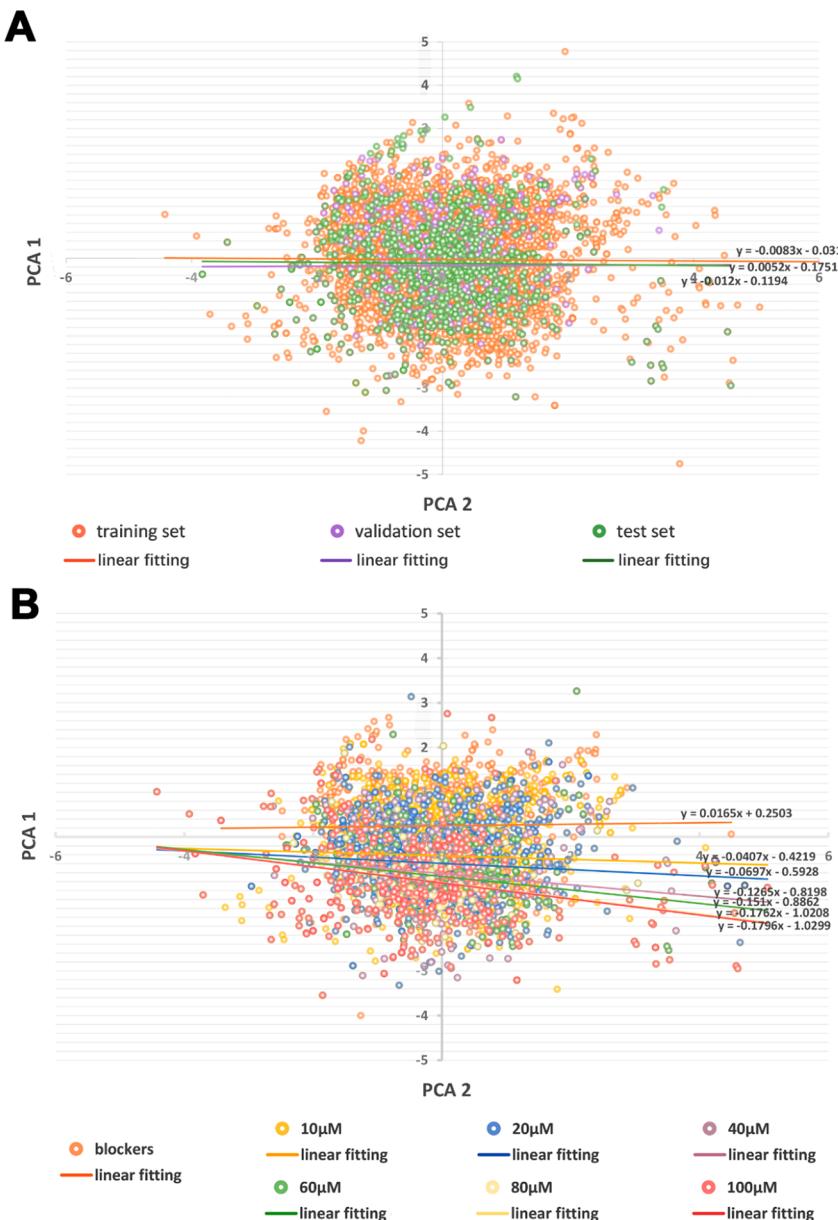


Figure 2. Distribution of chemical diversity. (A) Principal component analysis (PCA) of chemical diversity analysis across training sets, test sets, and validation sets. (B) PCA of chemical space between hERG blockers and hERG nonblockers. Linear fittings were utilized to compare the differentiation among different data sets. The closer the fitting lines they present, the more similarity of chemical space they have. The nonlinear fittings of chemical spaces for training set, test set, and validation set are provided in Supporting Information Figure S2.

large chemical space, to a certain extent, reflects the effectiveness of the applicability domain of a model.⁴⁹ In this study, the chemical space was analyzed using principal component analysis (PCA)⁵⁰ with 185 MOE descriptors as input. As demonstrated by the chemical space defined by the first two principal components in Figure 2A and Supporting Information Figure S2, high chemical diversity of data sets as well as a similar applicability domain⁵¹ among the compounds within training sets, test sets, and validation sets are observed. In addition, the PCA approach was also employed to observe the difference of chemical space between blocker compounds and decoys of different threshold values (Figure 2B). Obviously, as the value of decoy threshold increases, the distinction between blockers and decoys become more significant. Compared to the linear fitting line of blockers,

the fitting line of decoys with threshold values of 10 μ M shows highest parallelism, while 100 μ M exhibits the lowest.

Validation of Multitask DNN Models. The hyperparameter batch size defines the number of events that a model reads from the input at a time for stochasticity. In general, smaller batch size leads to a less accurate estimated gradient, but models trained with an oversize batch are more prone to be stuck in local optima.⁵² As discussed above, decoys from a higher threshold can be more distinct from the blockers but also lead to less samples. Insufficient sample size may bring unsatisfactory performance. Put together, we constructed various multitask DNN models with different batch sizes (32, 64, 128, 256, 512, and 1024) as well as various threshold values of decoy (10, 20, 40, 60, 80, and 100 μ M) for building the best model. We found that the best multitask DNN model was built with the batch size of 256, which

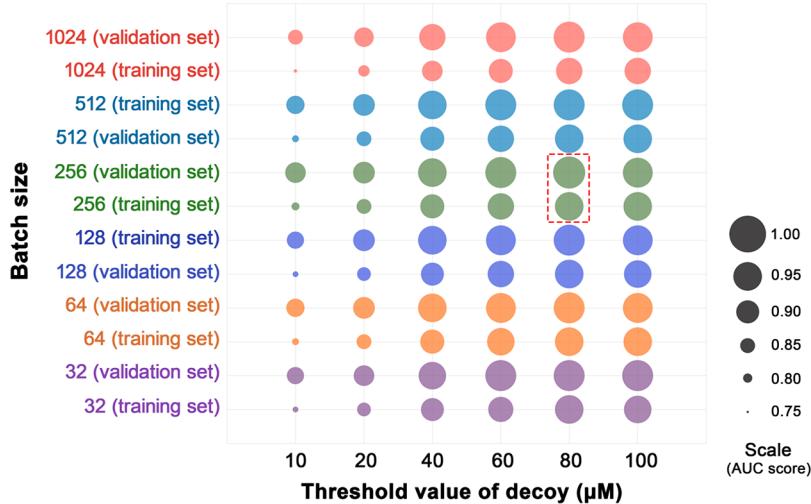


Figure 3. Comparison of the area under the receiver operating characteristic curve (AUC) values of multitask deep neural network (DNN) models. Multitask DNN were built by different batch sizes on training sets and validation sets. The size of circles denotes the value of AUC. The best results (the red dotted box) were achieved at the batch size of 256 and decoy threshold value of 80 μM .

Table 2. Area under the Receiver Operating Characteristic Curve (AUC) Values of Different Thresholds of Multitask DNN Models Constructed with Different Batch Sizes

batch size	training set						validation set					
	10 μM	20 μM	40 μM	60 μM	80 μM	100 μM	10 μM	20 μM	40 μM	60 μM	80 μM	100 μM
32	0.783	0.840	0.902	0.925	0.942	0.941	0.861	0.889	0.939	0.959	0.964	0.960
64	0.789	0.847	0.912	0.933	0.946	0.946	0.873	0.895	0.948	0.956	0.962	0.955
128	0.783	0.841	0.906	0.929	0.943	0.941	0.867	0.894	0.942	0.957	0.960	0.952
256	0.792	0.847	0.911	0.931	0.944	0.943	0.883	0.899	0.950	0.962	0.967	0.958
512	0.785	0.844	0.910	0.931	0.945	0.945	0.871	0.896	0.948	0.960	0.965	0.959
1,024	0.761	0.820	0.888	0.910	0.930	0.927	0.842	0.881	0.931	0.953	0.960	0.951

Table 3. Performance of the Best Multitask DNN Models (Batch Size = 256) for the Training Set and Validation Set^a

threshold value of decoy (μM)	training set						validation set					
	SE	SP	Q+	Q-	Q	AUC	SE	SP	Q+	Q-	Q	AUC
10	0.767	0.636	0.722	0.689	0.709	0.792	0.816	0.808	0.839	0.781	0.812	0.883
20	0.813	0.697	0.842	0.653	0.775	0.847	0.848	0.781	0.885	0.722	0.826	0.899
40	0.875	0.753	0.935	0.598	0.851	0.911	0.894	0.880	0.968	0.674	0.891	0.950
60	0.884	0.793	0.959	0.559	0.870	0.931	0.906	0.889	0.978	0.637	0.903	0.962
80	0.912	0.817	0.974	0.556	0.901	0.944	0.926	0.914	0.988	0.624	0.925	0.967
100	0.895	0.826	0.979	0.461	0.888	0.943	0.915	0.896	0.988	0.538	0.913	0.958

^aNote: SE: sensitivity; SP: specificity; Q+: prediction accuracy of hERG blockers; Q-: prediction accuracy of nonblockers; Q: overall predictive accuracy; AUC: the area under the receiver operating characteristic curve.

produced the best AUC value in most tasks (Figure 3). Detailed performance of multi-DNN models is listed in Table 2. As expected, the increasing threshold value of decoys does enhance the performance of model but stops at 80 μM , which may be a result of the lack of negative samples. To sum up, the model with batch size of 256 and decoy threshold value of 80 μM demonstrates greater predictive power than others by achieving the highest AUC values (0.967) and Q values (0.925) on the validation set (Table 3).

In this study, molecular properties and structural information were represented using MOE together with Mol2vec descriptors. To further clarify, multitask DNN models using the MOE descriptors only, Mol2vec descriptors only, and the combination of MOE and Mol2vec descriptors (MOE +Mol2vec) were developed independently with consistent hyperparameters and settings. We found that models building

on MOE+Mol2vec demonstrated a higher accuracy across all tasks on both training sets and validation sets (Figure 4). We thus selected the multitask DNN model building on MOE +Mol2vec descriptors with decoy threshold value of 80 μM as deephERG for further evaluation.

Comparison of Different Approaches. We further compared the performance of multitask DNN and several different approaches, including single-task DNN, GCNN, SVM, NB, and RF (Supporting Information, Table S5). We found that multitask DNN models offer better accuracies consistently over single-task DNN models across all tasks on both training set and validation set (Figure 5 and Supporting Information, Figure S3), indicating that multitask DNN can exploit representations learned across different tasks and can boost the performance on each task.

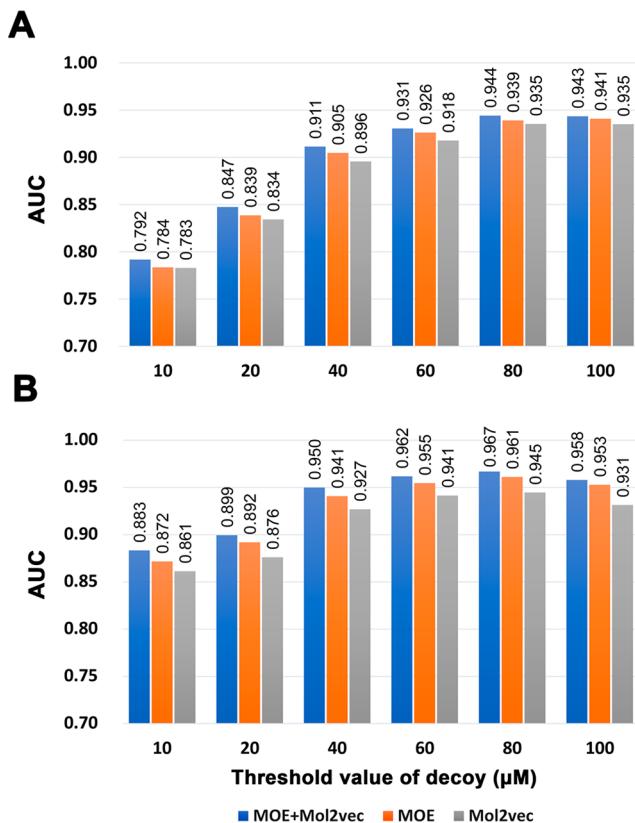


Figure 4. Comparison of the area under the receiver operating characteristic curve (AUC) values of multitask deep neural network (DNN) models on training set (A) and validation set (B). Multitask DNN models were built by different threshold values and three different combinations of descriptors: MOE+Mol2vec, Mol2vec alone, or MOE alone.

Comparison was also drawn between multitask DNN and three traditional machine learning methods (SVM, NB, and

RF). The performance of multitask DNN outperforms traditional machine learning methods (*Supporting Information, Table S5*). For example, as shown in Figure 5, the AUC value of 0.967 for multitask DNN on validation set with decoy threshold values of 80 μM is higher than those for NB (AUC = 0.922), SVM (AUC = 0.908), and RF (AUC = 0.950). In addition, we further compared deephERG with GCNN that is available from DeepChem.³⁷ We found that deephERG (AUC = 0.967) marginally outperforms that of GCNN (AUC = 0.959) on the validation set (*Supporting Information, Table S5*).

Identification of Potential hERG Blockers from Approved Drugs. Via deephERG, we next turned to perform risk assessment of hERG blockers for 1824 FDA approved small molecule drugs available from DrugBank database.³³ All drugs were processed by molecular washing and energy minimizing as described previously¹¹ after eliminating antibodies or ions drugs. Among 1,824 approved small molecule drugs, 539 drugs (29.6%) were computationally predicted to have potential hERG inhibitory activities ($IC_{50} \leq 10 \mu M$) by deephERG (Figure 6). Among 15 drug categories defined by Anatomical Therapeutic Chemical Classification System (ATC code), the five categories with the top number of predicted hERG blockers included: nervous system (N), cardiovascular system (C), alimentary tract and metabolism (A), antineoplastic and immunomodulating agents (L), and respiratory system (R). Some of our predictions are in accordance with literature evidence. For example, prochlorperazine, a typical antipsychotic agent, induced a concentration-dependent decrease in current amplitudes at the end of the voltage steps and tail currents of hERG.⁵³ Ivermectin, as a broad-spectrum antiparasitic agent, *in vitro* assay showed that it has an effect on the rapid delayed rectifier current mediated by the K(+) ion channel encoded by hERG.⁵⁴ In addition, a hERG inhibitory effect was observed on somatostatin and its derivatives as well.⁵⁵ The detailed prediction results for 1,824 approved drugs by deephERG are provided in *Supporting*

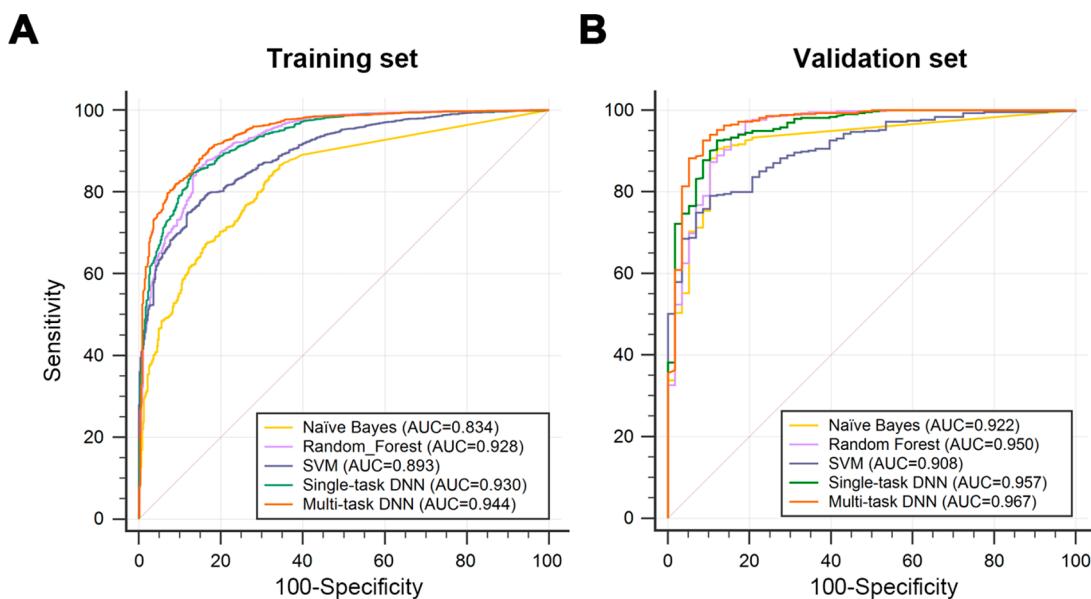


Figure 5. Receiver operating characteristic (ROC) curves of multitask deep neural network (DNN) model, single-task DNN model, and the three traditional machine learning methods: support vector machine (SVM), naïve Bayes, and random forest. All models were built and evaluated based on MOE+Mol2vec descriptors with biological end point of hERG at 80 μM on training set (A) and validation set (B). AUC: the area under the receiver operating characteristic curve.

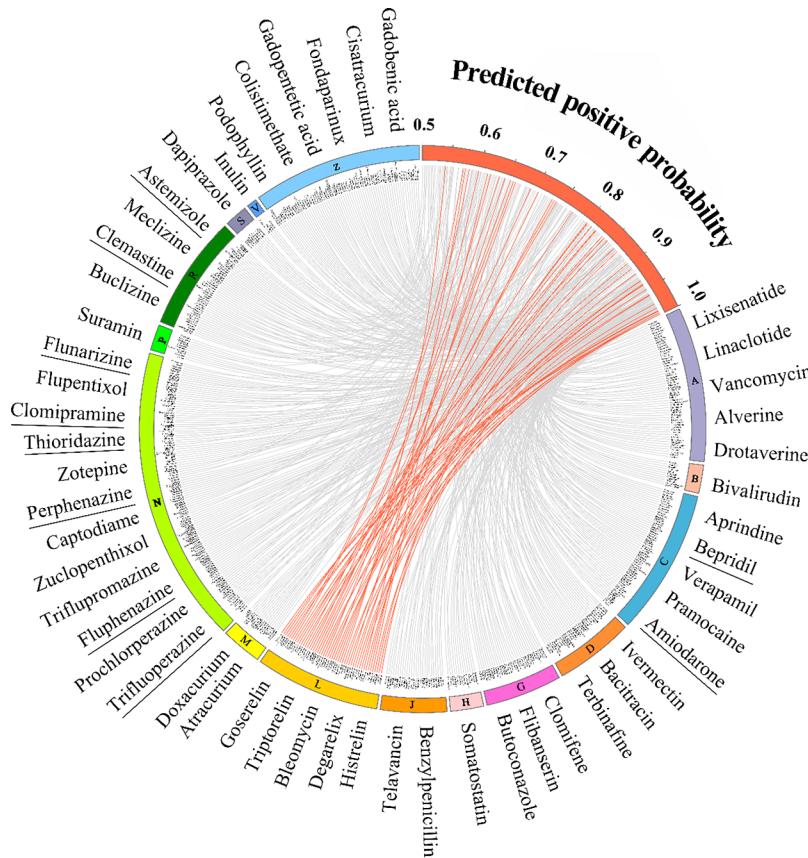


Figure 6. hERG-mediated cardiotoxic risk assessment for 1,824 FDA-approved drugs by deephERG. Note A: alimentary tract and metabolism; B: blood and blood forming organs; C: cardiovascular system; D: dermatologicals; G: systemic hormonal preparations; H: excluding sex hormones and insulins; J: antiinfectives for systemic use; L: antineoplastic and immunomodulating agents; M: musculo-skeletal system; N: nervous system; P: antiparasitic products, insecticides, and repellents; R: respiratory system; S: sensory organs; V: various; Z: unknown. The deephERG-predicted drugs with predicted positive probabilities more than 0.5 were illustrated. Drugs with the top 10% highest predicted positive probability in each category are displayed enlarged on the outer circle, while all the predicted positive drugs are presented on the inner circle. Drugs existing in the positives of training sets are underlined. The high-resolution version is provided in the Supporting Information.

Information, Table S6. We next turned to focus on the deephERG-predicted cardiotoxicity risk on approved anticancer agents (highlighted by red in Figure 6).

Evaluation of Cardiotoxicity Risk of Anticancer Agents by deephERG. The growing awareness of cardiovascular complications associated with cancer treatment has led to the emerging field of cardio-oncology (also known onco-cardiology), which centers on screening, monitoring, and treating cancer patients with cardiac dysfunction.^{56,57} Furthermore, it is also an exciting field because there is no guidelines and nonavailable FDA-approved therapeutics in terms of how to prevent new cardiotoxicity in cancer survivors. In addition to the myocardial cell, the hERG channel is also expressed highly in tumor cells and is related to the regulation of cell apoptosis and proliferation.^{58,59} Thus, we further investigated whether we could identify new hERG channel blockers on approved anticancer agents by deephERG (Supporting Information, Table S7). Figure 7 shows the more intuitive prediction results for 49 anticancer drugs by deephERG. In total, 15 out of 49 drugs whose hERG inhibitory activities have been validated by clinical case report or reported preclinical data are highlighted in bold (Supporting Information, Table S8).

Among protein kinase inhibitors, tyrosine kinase inhibitors (TKIs) occupy a large proportion. TKIs have played important roles in molecularly targeted treatment in multiple cancer

types. We found that 24 TKIs were predicted as hERG blockers by deephERG and 9 (recall = 37.5% [9/24]) of them had been validated by clinical case reports, *in vitro* assays, and the literature (Figure 7 and Supporting Information, Table S8). Among them, erlotinib was detected to have potential hERG inhibition at the concentration of 10 μ M by an automated patch-clamp assay, consistent with deephERG-prediction.⁶⁰ In addition, sunitinib, crizotinib, and nilotinib were also confirmed to block potently the hERG channel with IC₅₀ values of 0.5, 1.7, and 0.7 μ M, respectively.⁶⁰ Ponatinib, a multitargeted TKI and potent pan-ABL inhibitor, was temporarily withdrawn from the U.S. market due to severe vascular adverse events.⁶¹ Herein ponatinib was predicted as a hERG blocker by deephERG. An *in vitro* assay conducted in human embryonic kidney cells stably expressing the hERG potassium channel revealed that ponatinib inhibits hERG at a concentration above 1 μ M.⁶² Vandetanib, a multikinase inhibitor approved for treatment of multiple cancer types, has been reported to inhibit the hERG at 3 μ M detected by a whole-cell patch-clamp assay in transiently transfected HEK293 cells.⁶³ Finally, several new deephERG-predicted TKIs without literature data also need to be highly vigilant, such as bosutinib, gefitinib, afatinib, and dasatinib, which are warranted by experimental or clinical validation in the future.

In addition to TKIs, we also computationally predicted multiple cytotoxic chemotherapeutic agents as potential hERG

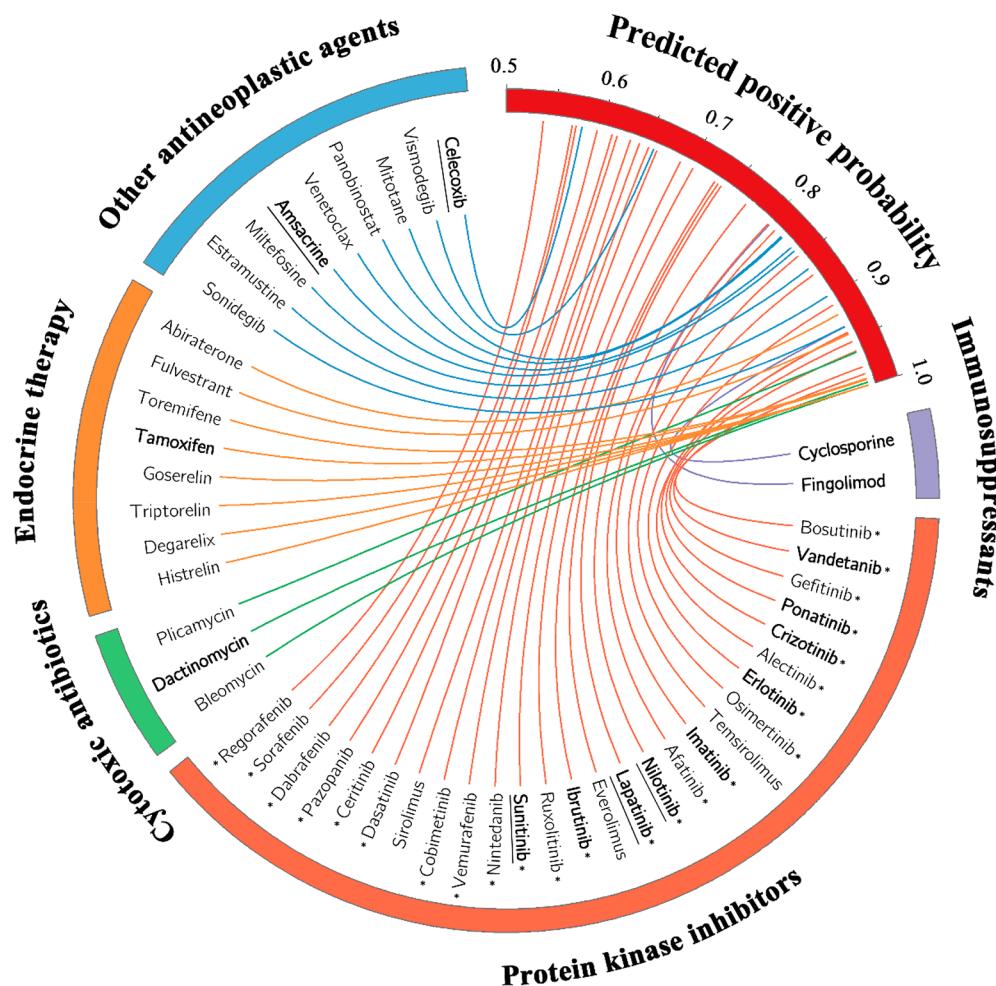


Figure 7. hERG-mediated cardiotoxic risk assessment for 49 approved antineoplastic drugs (including immunomodulating agents) by deephERG. Drugs which hERG inhibitory activity had been validated by reported experimental data are highlighted in bold font ([Supporting Information, Table S8](#)). Drugs existing in the positives of training sets are underlined, and tyrosine kinase inhibitors are marked with *. The deephERG-predicted drugs with predicted positive probabilities more than 0.5 were illustrated.

blockers as well by deephERG, such as dactinomycin, tamoxifen, and amsacrine. For example, the hERG inhibitory activity of dactinomycin from cytotoxic antibiotics and tamoxifen from endocrine therapy had been validated by *in vitro* assays.^{64,65} Tamoxifen inhibited hERG at -50 mV in a concentration-dependent manner with IC_{50} value of $1.2 \mu\text{M}$.⁶⁴ Dactinomycin markedly reduced the hERG mRNA levels as well.⁶⁵ Moreover, amsacrine, an antineoplastic agent used in acute lymphoblastic leukemia, had been confirmed to block hERG in HEK 293 cells with IC_{50} values of 209.4 nM .⁶⁶ Finally, we computationally identified multiple novel hERG blockers by deephERG, such as picamycin and bleomycin. In summary, 23% (10/44, excluding 5 duplicated drugs existing in the training set) of predictions of antineoplastic and immunomodulating agents had been validated successfully by clinical case reports, experiment evidence, and the literature, indicating that deephERG offers a useful tool for risk assessment of potential hERG-mediated cardiotoxicities in drug discovery and postmarketing surveillance.

■ DISCUSSION

In this study, we developed a deep learning approach, termed deephERG, for cardiotoxicity risk assessment of small molecules mediated by hERG blockers. We highlight several

improvements compared to previous studies.^{25,28,29,34} First, we assembled chemical diverse hERG channel blockers with well-defined experimental end points by our sizable efforts, which increase the applicability domain of deephERG models. To our knowledge, this is the largest publicly available data set with well-defined experimental hERG blocking bioactivity values. We believed that this data set is expected to represent comprehensive coverage of known hERG blockers. Second, it is undetermined to define compounds as hERG nonblockers, since which decoy threshold should be used to distinguish blockers and nonblockers has not yet reported previously. For example, in Wang's paper,²⁹ the threshold was $40 \mu\text{M}$, while Didzaiapetris³⁴ set it to $10 \mu\text{M}$ ([Supporting Information, Table S1](#)). In this study, we compared multiple decoy thresholds to search the most reasonable threshold ($IC_{50} = 80 \mu\text{M}$). Third, we showed that deephERG models built by multitask DNN algorithm outperformed single-task DNN, GCNN, and traditional machine learning models, including SVM, NB, and RF. This is consistent with previous studies that multitask DNN has better learning and adaptive ability compared to conventional machine learning approaches for drug discovery.^{67–71} Finally, all codes (multitask DNN and single-task DNN) and the Mol2vec approach used in this study can be conveniently implemented using a single python script, which

is freely available at <https://github.com/ChengF-Lab/deephERG>.

Several potential shortcomings should be acknowledged. First, although we assembled a large-scale compounds with well-defined experimental data on hERG channel from published literatures based on our sizable efforts, the incompleteness of samples still exists, especially for the negative compounds without known hERG inhibitory activities. As we discussed above, models trained based on higher threshold value of decoys had shown stronger predictive ability, but the insufficiency of learning samples hinder the advance. Therefore, a more comprehensive collection deriving from multiple sources should be further integrated in the future. Second, the Mol2vec approach applied for chemical representation in this study was based on a pretrained compound corpus of 19.9 million compounds, which may contain inefficient and redundant features. Replacement of the currently used corpus by a trimmed-down data set, such as marketed drugs, may characterize chemical structures more pointedly and efficiently. Third, as the hyperparameter selection of DNN is known to be very complicated, applying a new strategy⁷² to identify tunable hyperparameters may lead to better generalization ability of deephERG. Fourth, although the influence of model performance by different decoy threshold settings was explored, their underlying cause should be further delved. Fréchet ChemNet distance⁷³ is an innovative approach to compare the distribution of compounds, which may facilitate to explore the reason from the different evaluation matrices. Finally, although some of our predictions on approved drugs have already been confirmed by previously reported *in vitro* assay, clinical reports, and the literature, some novel deephERG-predicted hERG inhibitory drugs should be further validated by experimental assays or pharmacoepidemiologic analyses from real-world patient data in the future.^{11,74}

CONCLUSIONS

In this study, we proposed a deephERG framework to build predictive models for evaluation of hERG channel blockers based on the largest data set containing 7,889 compounds with well-defined experimental end points. DeephERG model built by a multitask DNN algorithm shows a satisfactory predictive ability with AUC values of 0.944 and 0.967 on training set and validation set respectively, which is superior than single-task DNN and traditional machine learning approaches, including support vector machine and naïve Bayes. Finally, we utilized the best deephERG model for potential hERG-mediated cardiotoxicity risk assessment on over 1,800 FDA-approved drugs. In case studies, we showed that the predicted cardiotoxicities for approved antineoplastic by deephERG were validated by reported experimental data from *in vitro* assays and clinical case reports. If broadly applied, deephERG presented here offer useful *in silico* tools for hERG-mediated cardiotoxicity risk assessment of small molecules in drug discovery and postmarketing surveillance.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jcim.8b00769](https://doi.org/10.1021/acs.jcim.8b00769).

High-resolution version of Figure 6. Detailed information on literature data sources of compounds with

experimental hERG blocking bioactivities value used in this study (Table S1), detailed description of literature-derived hERG blockers used in this study (Table S2), chemical information on hERG blockers and decoys (Table S3), chemical information on training set, test set, and validation set (Table S4), the area under the receiver operating characteristic curve (AUC) values of multitask deep neural network (DNN), single-task DNN, support vector machine, naïve Bayes, random forest, and graph convolutional neural network (GCNN) models across different decoy thresholds on training set and validation set (Table S5), detailed predictions for 1,824 FDA-approved drugs by deephERG (Table S6), detailed predictions for 49 approved antineoplastic drugs (including immunomodulating agents) by deephERG (Table S7), and detailed descriptions and PubMed ID (PMID) of the 15 deephERG-predicted antineoplastic drugs whose hERG inhibitory activities have been validated by clinical case report or reported preclinical data (Table S8). Average loss of multitask deep neural network (DNN) models based on different batch sizes (Figure S1), principal component analysis (PCA) of chemical diversity analysis across training sets, test sets, and validation sets with polynomial fitting curve (Figure S2), and comparison of the area under the receiver operating characteristic curve (AUC) value between multitask deep neural network (DNN) and single-task DNN models on training set and validation set (Figure S3) (ZIP)

AUTHOR INFORMATION

Corresponding Authors

*Email: fangjs@gzucm.edu.cn (J.F.).

*Email: f.cheng@neu.edu (F.C.). Phone: +1-216-4447654. Fax: +1-216-6361609.

ORCID

Chuiwu Cai: 0000-0002-6778-9697

Jiansong Fang: 0000-0002-6998-5384

Feixiong Cheng: 0000-0002-1736-2847

Notes

The authors declare no competing financial interest.

The code for deephERG is available at <https://github.com/ChengF-Lab/deephERG>.

ACKNOWLEDGMENTS

This work was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number K99HL138272 and R00HL138272 to F.C.

REFERENCES

- (1) Redfern, W. S.; Carlsson, L.; Davis, A. S.; Lynch, W. G.; Mackenzie, I.; Palethorpe, S.; Siegl, P. K. S.; Strang, I.; Sullivan, A. T.; Wallis, R. Relationships between Preclinical Cardiac Electrophysiology, Clinical QT Interval Prolongation and Torsade de Pointes for a Broad Range of Drugs: evidence for a provisional safety margin in drug development. *Cardiovasc. Res.* **2003**, *58*, 32–45.
- (2) Smith, P. L.; Baukowitz, T.; Yellen, G. The Inward Rectification Mechanism of the HERG Cardiac Potassium Channel. *Nature* **1996**, *379*, 833–836.
- (3) Sanguineti, M. C.; Jiang, C.; Curran, M. E.; Keating, M. T. A Mechanistic Link between an Inherited and an Acquired Cardiac Arrhythmia: HERG Encodes the IKr Potassium Channel. *Cell* **1995**, *81*, 299–307.

- (4) Villoutreix, B. O.; Tabouret, O. Computational Investigations of HERG Channel Blockers: New Insights and Current Predictive Models. *Adv. Drug Delivery Rev.* **2015**, *86*, 72–82.
- (5) Brown, A. M. Drugs, HERG and Sudden Death. *Cell Calcium* **2004**, *35*, 543–547.
- (6) Aronov, A. M. Predictive *in Silico* Modeling for HERG Channel Blockers. *Drug Discovery Today* **2005**, *10*, 149–155.
- (7) Witchel, H. J. The HERG Potassium Channel as a Therapeutic Target. *Expert Opin. Ther. Targets* **2007**, *11*, 321–336.
- (8) Raschi, E.; Vasina, V.; Poluzzi, E.; De Ponti, F. The HERG K⁺ Channel: Target and Antitarget Strategies in Drug Development. *Pharmacol. Res.* **2008**, *57*, 181–195.
- (9) Darpo, B.; Nebout, T.; Sager, P. T. Clinical Evaluation of QT/QTc Prolongation and Proarrhythmic Potential for Nonantiarrhythmic Drugs: the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use E14 guideline. *J. Clin. Pharmacol.* **2006**, *46*, 498–507.
- (10) Polak, S.; Wiśniowska, B.; Brandyś, J. Collation, Assessment and Analysis of Literature *in Vitro* Data on HERG Receptor Blocking Potency for Subsequent Modeling of Drugs' Cardiotoxic Properties. *J. Appl. Toxicol.* **2009**, *29*, 183–206.
- (11) Cai, C.; Fang, J.; Guo, P.; Wang, Q.; Hong, H.; Moslehi, J.; Cheng, F. *In Silico* Pharmacoepidemiologic Evaluation of Drug-Induced Cardiovascular Complications using Combined Classifiers. *J. Chem. Inf. Model.* **2018**, *58*, 943–956.
- (12) Lei, T.; Li, Y.; Song, Y.; Li, D.; Sun, H.; Hou, T. ADMET Evaluation in Drug Discovery: 15. Accurate Prediction of Rat Oral Acute Toxicity Using Relevance Vector Machine and Consensus Modeling. *J. Cheminf.* **2016**, *8*, 1–6.
- (13) Lei, T.; Chen, F.; Liu, H.; Sun, H.; Kang, Y.; Li, D.; Li, Y.; Hou, T. ADMET Evaluation in Drug Discovery. 17. Development of Quantitative and Qualitative Prediction Models for Chemical-Induced Respiratory Toxicity. *Mol. Pharmaceutics* **2017**, *14*, 2407–2421.
- (14) Lei, T.; Sun, H.; Kang, Y.; Zhu, F.; Liu, H.; Zhou, W.; Wang, Z.; Li, D.; Li, Y.; Hou, T. ADMET Evaluation in Drug Discovery. 18. Reliable Prediction of Chemical-Induced Urinary Tract Toxicity by Boosting Machine Learning Approaches. *Mol. Pharmaceutics* **2017**, *14*, 3935–3953.
- (15) Cheng, F.; Yu, Y.; Shen, J.; Yang, L.; Li, W.; Liu, G.; Lee, P. W.; Tang, Y. Classification of Cytochrome P450 Inhibitors and Non-inhibitors Using Combined Classifiers. *J. Chem. Inf. Model.* **2011**, *51*, 996–1011.
- (16) Cheng, F.; Lee, P. W.; et al. AdmetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *J. Chem. Inf. Model.* **2012**, *52*, 3099–3105.
- (17) Cheng, F.; Zhao, Z. Machine Learning-based Prediction of Drug-drug Interactions by Integrating Drug Phenotypic, Therapeutic, Chemical, and Genomic Properties. *J. Am. Med. Inform. Assn.* **2014**, *21*, 278–286.
- (18) Cheng, F.; Ikenaga, Y.; Zhou, Y.; Yu, Y.; Li, W.; Shen, J.; Du, Z.; Chen, L.; Xu, C.; Liu, G.; et al. In Silico Assessment of Chemical Biodegradability. *J. Chem. Inf. Model.* **2012**, *52*, 655–669.
- (19) Cheng, F.; et al. In Silico ADMET Prediction: Recent Advances, Current Challenges and Future trends. *Curr. Top. Med. Chem.* **2013**, *13*, 1273–1289.
- (20) Bains, W.; Basman, A.; White, C. HERG Binding Specificity and Binding Site Structure: Evidence from a Fragment-based Evolutionary Computing SAR Study. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 205–233.
- (21) Roche, O.; Trube, G.; Zuegge, J.; Pfleimlin, P.; Alanine, A.; Schneider, G. A Virtual Screening Method for Prediction of the hERG Potassium Channel Liability of Compound Libraries. *ChemBioChem* **2002**, *3*, 455–459.
- (22) Broccatelli, F.; Mannhold, R.; Moriconi, A.; Giuli, S.; Carosati, E. QSAR Modeling and Data Mining Link Torsades de Pointes Risk to the Interplay of Extent of Metabolism, Active Transport, and hERG Liability. *Mol. Pharmaceutics* **2012**, *9*, 2290–2301.
- (23) Sinha, N.; Sen, S. Predicting HERG Activities of Compounds from Their 3D Structures: Development and Evaluation of a Global Descriptors Based QSAR Model. *Eur. J. Med. Chem.* **2011**, *46*, 618–630.
- (24) Wang, S.; Li, Y.; Xu, L.; Li, D.; Hou, T. Recent Developments in Computational Prediction of HERG Blockage. *Curr. Top. Med. Chem.* **2013**, *13*, 1317–1326.
- (25) Chemi, G.; Gemma, S.; Campiani, G.; Brogi, S.; Butini, S.; Brindisi, M. Computational Tool for Fast *in Silico* Evaluation of hERG K⁺ Channel Affinity. *Front. Chem.* **2017**, *5*, 7.
- (26) Jing, Y.; Easter, A.; Peters, D.; Kim, N.; Enyedy, I. J. *In Silico* Prediction of hERG Inhibition. *Future Med. Chem.* **2015**, *7*, 571–586.
- (27) Wang, S.; Li, Y.; Wang, J.; Chen, L.; Zhang, L.; Yu, H.; Hou, T. ADMET Evaluation in Drug Discovery. 12. Development of Binary Classification Models for Prediction of hERG Potassium Channel Blockage. *Mol. Pharmaceutics* **2012**, *9*, 996–1010.
- (28) Doddareddy, M. R.; Klaasse, E. C.; Ijzerman, A. P.; Bender, A.; Shagupta. Prospective Validation of a Comprehensive *in Silico* hERG Model and its Applications to Commercial Compound and Drug Databases. *ChemMedChem* **2010**, *5*, 716–729.
- (29) Wang, S.; Sun, H.; Liu, H.; Li, D.; Li, Y.; Hou, T. ADMET Evaluation in Drug Discovery. 16. Predicting hERG Blockers by Combining Multiple Pharmacophores and Machine Learning Approaches. *Mol. Pharmaceutics* **2016**, *13*, 2855–2866.
- (30) Czodrowski, P. hERG me out. *J. Chem. Inf. Model.* **2013**, *53*, 2240–2251.
- (31) Vilar, S.; Cozza, G.; Moro, S. Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Curr. Top. Med. Chem.* **2008**, *8*, 1555–1572.
- (32) Jaeger, S.; Fulle, S.; Turk, S. Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition. *J. Chem. Inf. Model.* **2018**, *58*, 27–35.
- (33) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; et al. DrugBank 4.0: Shedding New Light on Drug Metabolism. *Nucleic Acids Res.* **2014**, *42*, D1091–1097.
- (34) Didziapetris, R.; Lanevskij, K. Compilation and Physicochemical Classification Analysis of a Diverse HERG Inhibition Database. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 1175–1188.
- (35) Feinberg, E. N.; Sur, D.; Husic, B. E.; Mai, D.; Li, Y.; Yang, J.; Ramsundar, B.; Pande, V. S. Spatial Graph Convolutions for Drug Discovery. *arXiv.org* **2014**, No. 1803.04465.
- (36) Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv.org* **2013**, No. 1301.3781.
- (37) Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.; Pande, V. MoleculeNet: A Benchmark for Molecular Machine Learning. *Chem. Sci.* **2018**, *9*, 513–530.
- (38) Demšar, J.; Curk, T.; Erjavec, A.; Goru, C.; Hočevar, T.; Milutinović, M.; Možina, M.; Polajnar, M.; Toplak, M.; Starič, A. Orange: Data Mining Toolbox in Python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
- (39) Simm, J.; Klambauer, G.; Arany, A.; Steijaert, M.; Wegner, J. K.; Gustin, E.; Chupakhin, V.; Chong, Y. T.; Vialard, J.; Buijnsters, P.; et al. Repurposing High-Throughput Image Assays Enables Biological Activity Prediction for Drug Discovery. *Cell Chem. Biol.* **2018**, *25*, 611–618.
- (40) Kingma, D.; Ba, J. Adam: A Method for StoCchastic Optimization. *arXiv.org* **2014**, No. 1412.6980.
- (41) Hahnloser, R. H.; Sarpeshkar, R.; Mahowald, M. A.; Douglas, R. J.; Seung, H. S. Digital Selection and Analogue Amplification Coexist in a Cortex-inspired Silicon Circuit. *Nature* **2000**, *405*, 947–951.
- (42) Bergstra, J.; Bengio, Y. Random Search for Hyper-parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
- (43) Ramsundar, B.; Liu, B.; Wu, Z.; Verras, A.; Tudor, M.; Sheridan, R. P.; Pande, V. S. Is Multitask Deep Learning Practical for Pharma? *J. Chem. Inf. Model.* **2017**, *57*, 2068–2076.
- (44) Byvatov, E.; Schneider, G. Support Vector Machine Applications in Bioinformatics. *Appl. Bioinf.* **2003**, *2*, 67–77.

- (45) Fang, J.; Yang, R.; Gao, L.; Zhou, D.; Yang, S.; Liu, A.-l.; Du, G.-h. Predictions of BuchE Inhibitors Using Support Vector Machine(SVM) and Naive Bayesian Classification Techniques. *J. Chem. Inf. Model.* **2013**, *53*, 3009–3020.
- (46) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.
- (47) Chang, C. C.; Lin, C. J. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 389–396.
- (48) Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874.
- (49) Cai, C.; Wu, Q.; Luo, Y.; Ma, H.; Shen, J.; Zhang, Y.; Yang, L.; Chen, Y.; Wen, Z.; Wang, Q. *In Silico* Prediction of ROCK II Inhibitors by Different Classification Approaches. *Mol. Diversity* **2017**, *21*, 791–807.
- (50) Ma, S.; Dai, Y. Principal Component Analysis Based Methods in Bioinformatics Studies. *Briefings Bioinf.* **2011**, *12*, 714–722.
- (51) Weaver, S.; Gleeson, M. P. The Importance of the Domain of Applicability in QSAR Modeling. *J. Mol. Graphics Modell.* **2008**, *26*, 1315–1326.
- (52) Keskar, N. S.; Mudigere, D.; Nocedal, J.; Smelyanskiy, M.; Tang, P. T. P. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. *arXiv.org* **2016**, No. 1609.04836.
- (53) Kim, M. D.; Eun, S. Y.; Jo, S. H. Blockade of HERG Human K⁺ Channel and IKR of Guinea Pig Cardiomyocytes by Prochlorperazine. *Eur. J. Pharmacol.* **2006**, *544*, 82–90.
- (54) Kauthale, R. R.; Dadarkar, S. S.; Husain, R.; Karande, V. V.; Gatne, M. M. Assessment of Temperature-Induced HERG Channel Blockade Variation by Drugs. *J. Appl. Toxicol.* **2015**, *35*, 799–805.
- (55) Yamasaki, T.; Hirose, H.; Yamashita, T.; Takakura, N.; Morimoto, S.; Nakahata, T.; Kina, A.; Nakano, Y.; Okano, T. Y.; Sugama, J.; et al. Discovery of Novel Somatostatin Receptor Subtype 5 (SSTR5) Antagonists: Pharmacological Studies and Design to Improve Pharmacokinetic Profiles and human Ether-a-go-go-related Gene (hERG) Inhibition. *Bioorg. Med. Chem.* **2017**, *25*, 4153–4162.
- (56) Moslehi, J. J. Cardiovascular Toxic Effects of Targeted Cancer Therapies. *N. Engl. J. Med.* **2016**, *375*, 1457–1467.
- (57) Sharma, A.; Burridge, P. W.; McKeithan, W. L.; Serrano, R.; Shukla, P.; Sayed, N.; Churko, J. M.; Kitani, T.; Wu, H.; Holmström, A.; et al. High-throughput Screening of Tyrosine Kinase Inhibitor Cardiotoxicity with Human Induced Pluripotent Stem Cells. *Sci. Transl. Med.* **2017**, *9*, eaaf2584.
- (58) Jehle, J.; Schweizer, P. A.; Katus, H. A.; Thomas, D. Novel Roles for HERG K⁺ Channels in Cell Proliferation and Apoptosis. *Cell Death Dis.* **2011**, *2*, e193–200.
- (59) Wang, J.; Zhang, Y.; Cao, L.; Han, H.; Yang, B.; Nattel, S.; Wang, Z.; Wang, J. HERG K⁺ Channel, a Regulator of Tumor Cell Apoptosis and Proliferation. *Cancer Res.* **2002**, *62*, 4843–4848.
- (60) Doherty, K. R.; Wappel, R. L.; Talbert, D. R.; Trusk, P. B.; Moran, D. M.; Kramer, J. W.; Brown, A. M.; Shell, S. A.; Bacus, S. Multi-parameter *in Vitro* Toxicity Testing of Crizotinib, Sunitinib, Erlotinib, and Nilotinib in Human Cardiomyocytes. *Toxicol. Appl. Pharmacol.* **2013**, *272*, 245–255.
- (61) Talbert, D. R.; Doherty, K. R.; Trusk, P. B.; Moran, D. M.; Shell, S. A.; Bacus, S. A Multi-parameter *in Vitro* Screen in Human Stem Cell-derived Cardiomyocytes Identifies Ponatinib-induced Structural and Functional Cardiac Toxicity. *Toxicol. Sci.* **2015**, *143*, 147–155.
- (62) Sonnichsen, D.; Dorer, D. J.; Cortes, J.; Talpaz, M.; Deininger, M. W.; Shah, N. P.; Kantarjian, H. M.; Bixby, D.; Mauro, M. J.; Flinn, I. W.; et al. Analysis of the Potential Effect of Ponatinib on the QTc Interval in Patients with Refractory Hematological Malignancies. *Cancer Chemother. Pharmacol.* **2013**, *71*, 1599–1607.
- (63) Lee, H. A.; Hyun, S. A.; Byun, B.; Chae, J. H.; Kim, K. S. Electrophysiological Mechanisms of Vandetanib-induced Cardiotoxicity: Comparison of Action Potentials in Rabbit Purkinje Fibers and Pluripotent Stem Cell-derived Cardiomyocytes. *PLoS One* **2018**, *13*, e0195577.
- (64) Chae, Y. J.; Lee, K. J.; Lee, H. J.; Sung, K. W.; Choi, J. S.; Lee, E. H.; Hahn, S. J. Endoxifen, the Active Metabolite of Tamoxifen, Inhibits Cloned HERG Potassium Channels. *Eur. J. Pharmacol.* **2015**, *752*, 1–7.
- (65) Krishnan, Y.; Li, Y.; Zheng, R.; Kanda, V.; McDonald, T. V. Mechanisms Underlying the Protein-kinase Mediated Regulation of the HERG Potassium Channel Synthesis. *Biochim. Biophys. Acta, Mol. Cell Res.* **2012**, *1823*, 1273–1284.
- (66) Thomas, D.; Hammerling, B. C.; Wu, K.; Wimmer, A. B.; Ficker, E. K.; Kirsch, G. E.; Kochan, M. C.; Wible, B. A.; Scholz, E. P.; Zitron, E.; et al. Inhibition of Cardiac HERG Currents by the DNA Topoisomerase II Inhibitor Amsacrine: Mode of Action. *Br. J. Pharmacol.* **2004**, *142*, 485–494.
- (67) Dahl, G. E.; Jaitly, N.; Salakhutdinov, R. Multi-task Neural Networks for QSAR Predictions. *arXiv.org* **2014**, No. 1406.1231.
- (68) Ramsundar, B.; Kearnes, S.; Riley, P.; Webster, D.; Konerding, D.; Pande, V. Massively Multitask Networks for Drug Discovery. *arXiv.org* **2015**, No. 1502.02072.
- (69) Ma, J.; Sheridan, R. P.; Liaw, A.; Dahl, G.; Svetnik, V. Deep Neural Nets as a Method for Quantitative Structure-Activity Relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263–274.
- (70) Li, X.; Xu, Y.; Lai, L.; Pei, J. Prediction of Human Cytochrome P450 Inhibition Using a Multitask Deep Autoencoder Neural Network. *Mol. Pharmaceutics* **2018**, *15*, 4336–4345.
- (71) Mayr, A.; Klambauer, G.; Unterthiner, T.; Steijaert, M.; Wegner, J. K.; Ceulemans, H.; Clevert, D. A.; Hochreiter, S. Large-scale Comparison of Machine Learning Methods for Drug Target Prediction on ChEMBL. *Chem. Sci.* **2018**, *9*, 5441–5451.
- (72) Zhou, Y.; Cahya, S.; Combs, A. S.; Nicolaou, C.; Wang, J.; Desai, V. P.; Shen, J. Exploring Tunable Hyperparameters for Deep Neural Network with Industrial ADME Data Sets. *J. Chem. Inf. Model.* **2019**, DOI: [10.1021/acs.jcim.8b00671](https://doi.org/10.1021/acs.jcim.8b00671).
- (73) Preuer, K.; Renz, P.; Unterthiner, T.; Hochreiter, S.; Klambauer, G. Fréchet ChemNet Distance: A Metric for Generative Models for Molecules in Drug Discovery. *J. Chem. Inf. Model.* **2018**, *58*, 1736–1741.
- (74) Cheng, F.; Desai, R. J.; Handy, D. E.; Wang, R.; Schneeweiss, S.; Barabási, A.; Loscalzo, J. Network-based Approach to Prediction and Population-based Validation of *in Silico* Drug Repurposing. *Nat. Commun.* **2018**, *9*, 2691.