A new classifier-based strategy for *in-silico* ion-channel cardiac drug safety assessment

Hitesh B. Mistry[1]*, Mark R. Davies[2] and Giovanni Y. Di Veroli[3]

[1] Manchester Pharmacy School, University of Manchester, Manchester, UK, [2] QT-Informatics Limited, Macclesfield, UK,
[3] Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK

There is currently a strong interest in using high-throughput *in-vitro* ion-channel screening data to make predictions regarding the cardiac toxicity potential of a new compound in both animal and human studies. A recent FDA think tank encourages the use of biophysical mathematical models of cardiac myocytes for this prediction task. However, it remains unclear whether this approach is the most appropriate. Here we examine five literature data-sets that have been used to support the use of four different biophysical models and one statistical model for predicting cardiac toxicity in numerous species using various endpoints. We propose a simple model that represents the balance between repolarisation and depolarisation forces and compare the predictive power of the model against the original results (leave-one-out cross-validation). Our model showed equivalent performance when compared to the four biophysical models and one statistical model. We therefore conclude that this approach should be further investigated in the context of early cardiac safety screening when *in-vitro* potency data is generated.

Keywords: cardiac toxicity, mathematical model, ion-channel pharmacology, predictive pharmacology, biostatistics

Introduction

The consequence of a drug carrying cardiac liability is a significant cause of concern to the Pharmaceutical industry. In the late 1980s to the early 1990s, a number of significant withdrawals from market occurred due to the detection of drug induced cardiac arrhythmias, e.g., Torsades de Pointes (ventricular tachycardia) such as prenylamine, lidoflazine, and terodiline from UK and EU market (Stockbridge et al., 2013). The exact number of drugs that cause Torsades de Pointes is unknown, however the CredibleMeds initiative (www.crediblemeds.org) has cataloged 42 drugs with a "Known Risk of Torsades de Pointes." These arrhythmias were mostly attributed to drugs causing significant prolongation to the QTc interval (time taken for ventricular depolarisation and repolaristaion), which is in turn related to the delay in repolarisation of the ventricular wall. Causative link was then established with drug blocking of the hERG (*human-ether-a-go-go*) ion-channel within ventricular myocytes (Pollard et al., 2010). As a consequence, routine screening of compounds via *in-vitro* high-throughput screening (HTS) devices was quickly incorporated into early stages of drug-development (Pollard et al., 2010). Initial screens focused only on the hERG channel but in recent years it has become apparent that other ion-channels might critically affect cardiac electrophysiology. In particular, hCav1.2 and hNav1.5 have been recognized as key depolarising ion channels with important roles in the mechanisms causing arrhythmia,

e.g., long QT syndromes LQT3 and LQT8 syndromes (Lehnart et al., 2007). Therefore, screenings have now also been extended to include these other ion-channels (Cavero and Holzgrefe, 2014).

Research efforts have highlighted the potential use of *in-silico* biophysical models (Trayanova, 2011) of cardiac myocytes to predict the cardiac risk *in-vivo* or even in the clinical setting, based on *in-vitro* ion-channel screening data (Cavero and Holzgrefe, 2014). These models describe the dynamic opening and closing of ion-channels and resulting temporal variation of cell Action Potential (AP) via a set of differential equations. They are parameterized based on experimental data from electrophysiological recordings of isolated ion-channels and also whole cell AP recordings. Their goal is generally to produce a descriptive model of the cardiac myocyte which is then used to better understand general cardiac biology. Nevertheless, there is still uncertainty as to which model is best suited to assist in cardiotoxicity prediction.

Answering these questions is in line with recent initiatives from a FDA sponsored think tank suggesting the usage of *in-silico* tools in correlating non-clinical *in-vitro* studies with proarrhythmic risk (Sager et al., 2014). Two techniques have been used in the literature: (1) biophysical models which describe the dynamics of a cardiac myocyte through differential equations of which there are 4 examples (Mirams et al., 2011, 2014; Davies et al., 2012; Beattie et al., 2013) and (2) statistical models which focus on known ion-channel pharmacology of which there is only one literature example (Kramer et al., 2013). Here we investigate an alternative strategy based on a one-equation classifier model and show that similar predictive power can be obtained with this strategy. In particular, we highlight the capacity of such a model in handling all datasets, in contrast with the original studies where a specific model was used at each time. The models predictive power within each data-set under consideration is also assessed via a leave-one-out cross validation exercise where an optimal parameter set for each data-set is used. We will then discuss the advantages of this alternative approach.

Materials and Methods

Data-Sets

All data-sets are reported in the Supplementary Material. Here we present a brief summary:

- **Human 1** (Kramer et al., 2013) contains 55 compounds and assessed the Torsades de Pointes risk of each compound. All ion-channel potency data was generated from two *in-vitro* HTS platforms, Qpatch, and PatchXpress. The cell lines used were HEK293 and CHO (Chinese Hamster Ovary). The *in-silico* model used within that study was a statistical (logistic regression) model which classified a compound as posing a Torsades de Pointes risk or not.
- **Human 2** (Mirams et al., 2011) contains 31 compounds and assessed the Torsades de Pointes risk of each compound. Ion-channel potency data was derived from numerous literature *in-vitro* reports. The *in-silico* model used within that study was a biophysical model (39 differential equations) which classified compounds into one of four Torsades de Pointes risk

categories. This was then simplified to a binary classification question of whether a compound posed a Torsades de Pointes risk or not.

- **Human 3** (Mirams et al., 2014) contains 34 compounds and assessed the QTc prolongation potential of each compound. We investigated the data-set which gave the authors the best result. This contained *in-vitro* hERG manual patch-clamp data (obtained from regulatory documents for each compound) and *in-vitro* HTS data for the other ion-channels (IonWorks Quattro using CHO cells). The *in-silico* model used within that study was a biophysical model (41 differential equations) which classified compounds as QTc prolongers or not at the mean drug concentration. QTc prolongation was defined by a mean increase in QTc interval of at least 5 ms.
- **Dog** (Davies et al., 2012) contains 53 compounds and assessed the time taken for the canine AP duration to repolarise by 90% (APD_{90}) within the canine AP *in-vitro* assay. All ion-channel potency data was generated from the same *in-vitro* HTS platform, IonWorks\FLIPR with CHO cells. The *in-silico* model used within that study was a biophysical model (29 differential equations) which classified compounds to one of three categories: prolongation, shortening or no effect. Prolongation/shortening was defined by a greater than 10% change in either direction compared to control at a given drug concentration.
- **Rabbit** (Beattie et al., 2013) contains numerous data-sets with potency estimates generated from different platforms (*in-vitro* and *in-silico* i.e., computational chemistry predicted potency estimates) and assessed the change in QTc within the *ex-vivo* rabbit ventricular wedge assay. Here we considered the dataset that gave the authors the best result there. This contained a total of 77 compounds where all the ion-channel potency data was generated from the same *in-vitro* HTS platform, PatchXpress with HEK293 and CHO cells. The *in-silico* model used within that study was a biophysical model (45 differential equations) which classified compounds in the same way as the above dog study.

Drug Effect

The ion-channel screening data extracted from the articles was converted into block or agonism of an ion-channel at a given drug concentration of interest in the following way. We first defined the effect S of a drug on ion-channel x as:

$$S_x = 1 - \frac{1}{1 + \left(\frac{IC_{50}}{[D]}\right)^n},$$

for an inhibitor and

$$S_x = 1 + \frac{E_{max}}{1 + \left(\frac{EC_{50}}{[D]}\right)^n},$$

for an agonist, where $x = I_{Kr}$, $x = I_{CaL}$ or $x = I_{Na}$ refer to the hERG, hCav1.2, and hNav 1.5 ion-channels respectively, $[D]$ is the concentration of drug, IC_{50} is the amount of drug required to reduce the effect by 50%, EC_{50} is the amount of drug required

to increase the effect by $E_{max}/2$, E_{max} the maximum effect of an agonist and n the hill coefficient. The IC_{50}, EC_{50}, E_{max}, and n values were taken as per the published data-sets. There, drug effects within bio-physical models were modeled by scaling conductances of ion-channels by S_x which was also done here. Only effects against the above three ion-channels were considered as they were consistently measured across all studies.

One-Equation Classifier Model

It is well understood that the hERG channel is responsible for repolarisation and that hCav1.2 and hNav1.5 ion-channels are largely responsible for depolarisation of the action potential. We hypothesized that the balance between the block/agonism of depolarisation and repolarisation forces is the main parameter driving drug effects' predictions. The following classifier was suggested:

$$Z = \frac{1 + a_0 S_{CaL} + a_1 S_{Na}}{1 + a_2 S_{Kr}}$$

where a_0, a_1, and a_2 are coefficients that need to be determined. For simplicity, we initially chose to set the coefficients in the model to 1, as the first goal of this study was to assess the discriminatory value of the model without specific parameters optimization.

ROC Curve

The model was initially assessed through a receiver operating characteristic (ROC) curve analysis (Zou et al., 2007). The ROC curve is a plot of the sensitivity (true positive rate) as a function of 1—specificity (i.e., false positive rate) for different cut-off values of a discriminant variable. The ROC curve can be assessed by looking at the obtained area under the curve (AUC). A value equal to 1 implies 100% sensitivity/specificity while a value equal to 0.5 represents 50% sensitivity/specificity. The ROC curve can also be inspected visually by comparing it to the line of unity: the further the curve is above this line the better.

Within this analysis, the following binary classification questions were asked:

(1) For the dog and rabbit data-sets:

  (a) Does a compound cause prolongation?
  (b) Does a compound cause shortening?

(2) For human 1 and 2 data-sets, does a compound have a Torsades de Pointes risk or not [definitions as per original articles (Mirams et al., 2011; Kramer et al., 2013)]?

(3) For human 3 data-set, does a compound cause QTc prolongation?

For the animal data-sets the above setup does mean a compound can be classified as both prolongation and shortening, however this is relaxed when considering the predictivity of the model as described below.

Leave-One-Out Cross Validation

After assessing the structural models discriminatory value through a ROC analysis we assessed the predictive power of the

model via a leave-one-out cross validation. We used $n − 1$ compounds to train the model and the nth compound to test it. This process was repeated until all n compounds were used as a test-set. Here the model parameter values, a_0, a_1, and a_2, were optimized for each data-set (see Supplementary Material). Compounds within the animal data-sets were classified as causing prolongation, no-effect or shortening. The result of this analysis was summarized through sensitivity, specificity and balanced accuracy [i.e., (sensitivity + specificity)/2] metrics. These metrics were then compared to the reported results of the original articles model predictions.
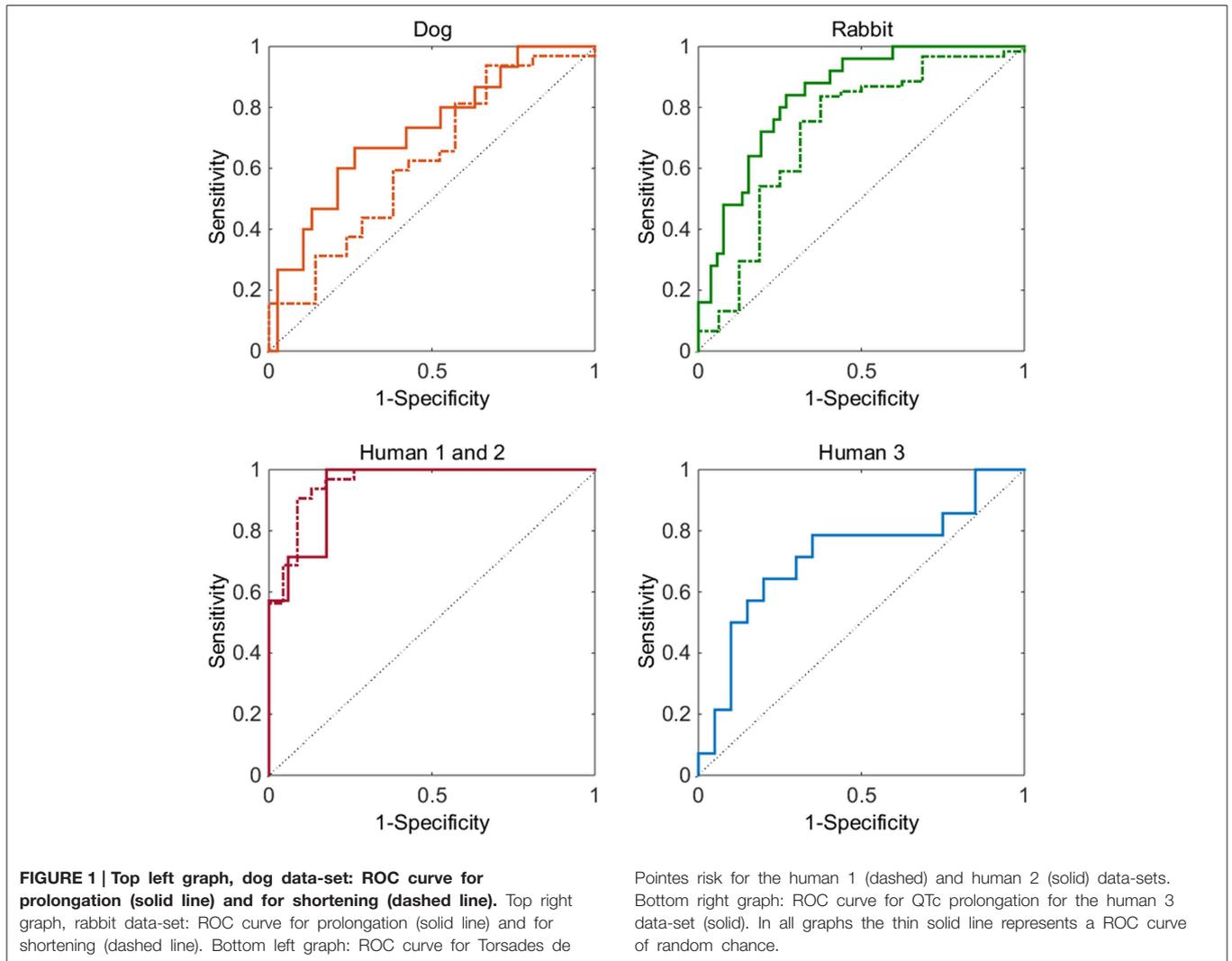
Results

ROC Analysis and Model Comparison

Overall it can be seen (**Figure 1**) that the ROC curve for the five cases investigated sit comfortably above the identity line. This shows that the model demonstrates good discriminatory power for all datasets without data-set specific parameter optimization. The strongest result can be seen for assessing whether a compound has a Torsades de Pointes risk or not (bottom left graph of **Figure 1**). The AUC values of the ROC curves for human 1 (dashed line) and 2 (solid line) data-sets are 0.94 and 0.96, respectively. These values are considered to be excellent and are close to a perfect classifier. The signal for QTc prolongation was also found to be good (ROC AUC of 0.72). The model also performed well for the animal data-sets: the ROC AUC values for dog and rabbit prolongation were 0.72 and 0.84 respectively, while the AUC values for shortening were 0.63 and 0.72 respectively.

We then performed a leave-one-out cross validation where model parameters were allowed to vary and an optimum set derived for each data-set. The results of this additional analysis and comparison of our model's prediction to the reported predictions can be seen in **Table 1**. Our model results were overall comparable to the published models' results. Sensitivity was increased in all data-sets except the Human 1 and 2 data-sets where the difference was minimal. Specificity was in good agreement against all the literature models. The balanced accuracy was either equal or improved over the current published models except in the Human 2 data-set where our model performed slightly worse than the biophysical model used there.

Discussion

There is currently strong interest from both regulatory agencies and pharmaceutical companies in the use of *in-silico* models of cardiac myocytes as a tool for integrating the cardiac ion-channel liability of new compounds. This interest is motivated by the question of predicting the effects of a new compound's cardiac risk in animal or man simply by using *in-vitro* ion-channel potency data (Cavero and Holzgrefe, 2014). This question has been predominantly answered to date by either empirically defining cut-off thresholds for IC_{50} values or by using *in-silico* approaches with biophysical models (Mirams et al., 2011, 2014; Davies et al., 2012; Beattie et al., 2013). Only one statistical model approach have been used to date to our

FIGURE 1 | Top left graph, dog data-set: ROC curve for prolongation (solid line) and for shortening (dashed line). Top right graph, rabbit data-set: ROC curve for prolongation (solid line) and for shortening (dashed line). Bottom left graph: ROC curve for Torsades de Pointes risk for the human 1 (dashed) and human 2 (solid) data-sets. Bottom right graph: ROC curve for QTc prolongation for the human 3 data-set (solid). In all graphs the thin solid line represents a ROC curve of random chance.

TABLE 1 | Sensitivity, specificity, and balanced accuracy of our model compared with the original models' predictions.

Data-set Binary Test Question (Yes/No)	Sensitivity%		Specificity%		Balanced Accuracy%	
	This model	Original model	This model	Original model	This model	Original model
Human 1 Torsade de Pointes Risk	97	97	91	83	94	90
Human 2 Torsade de Pointes Risks	93	100	97	97	95	98.5
Human 3 QTc Prolongation	57	50	90	85	73.5	67.5
Rabbit Prolongation	96	72	73	81	84.5	66.5
Rabbit Shortening	50	31	87	80	67.5	60
Dog Prolongation	67	60	84	76	75.5	68
Dog Shortening	67	48	78	87	72.5	67.5

Better metrics are highlighted in green. Worse metrics are highlighted in red.

knowledge (Kramer et al., 2013). Biophysical models benefit by incorporating representations of cardiac myocyte biology that may provide an intuitive leap for the experienced electrophysiologist to articulate depolarization mechanisms. Indeed, there are now many models representing different cell types and species available (see for instance the CellML database http://models.cellml.org/cellml).

Nevertheless, these biophysical models can suffer from their size and display complex behavior which can be highly parameter-sensitive (Gutenkunst et al., 2007). Their formulations vary greatly between species and cell-type and there are also competing formulations for the same species and cell-type. It is difficult to fully validate such models in practice. Moreover, this pre-existing complexity can be compounded by the additional incorporation of other major biological processes such as drug binding kinetics, regulatory signaling or spatio-temporal effects as well as statistical assessment. Thus, we wanted to explore how a model that minimizes mathematical formulation to describe AP prolongation compares to biophysical models in its predictive capacity. Such a tool, which is far less complex than current biophysical models, could facilitate the additional incorporation of other fundamental processes (Lu et al., 2013) and of statistical considerations (Chain et al., 2011).

We introduced a new classifier-type model, which is relatively simple (one equation) and can be directly interpreted in terms of the balance between depolarisation and repolarisation forces. The model performance was assessed via a ROC analysis and a leave-one-out cross validation across various literature data-sets. When comparing this model performance against published biophysical models, we found comparable results (same scoring system as per original studies was used). As a consequence, these results suggest that biophysical models might not provide significant benefit over the approach taken here to describe the correlation between ion-channel block and compound risk. It is possible that these models may be better suited to understand and articulate the cellular and tissue processes mechanistically, without having necessarily reached the required maturity that would imply added benefit in terms of predictive pharmacology.

With our new approach, other biological processes can be also incorporated by emphasizing on their influence rather than their precise mathematical description. The influence of other ion channels such as I_{k1}, I_{NaCa}, I_{to} or I_{Ks} can be incorporated in similar ways as per the ion-channels considered here. We can also imagine the possibility of accounting for slow vs. fast ion-channel block with additional parameters. New approaches have indeed emerged recently to elucidate drug binding dynamics using modern HTS assays (Di Veroli et al., 2013, 2014). Other processes such as spatio-temporal effects or regulatory processes can be integrated as separated factors and could also account for variability across populations. Furthermore, a simple model like this is easier to implement for scientists of all backgrounds and does not require extensive amount of domain specific knowledge. It can potentially lead to both cost and time-savings for safety pharmacology groups who are interested in assessing ion-channel related cardiac toxicity risk of new compounds.

All mathematical models, either simple or complex, might be followed later on by experiments to test for cardiac safety screening. Nevertheless, they represent a valuable tool for assessing risk and prioritizing new compounds at early stages. Biophysical models explicitly describe the mechanisms behind AP propagation and other important biological processes and are thus invaluable for gaining qualitative understanding of biological processes. Although our simple model does not claim to provide the level of description given by biophysical models, our results showed similar predictive power when correlating ion-channel block with AP or QT prolongation/shortening using published datasets. At the same time, our model was applied across different species and also avoided model complexity and its inherent risks. Thus, we propose this approach to form the basis of data-driven systems pharmacology strategies as it simplifies the incorporation of other aspects such as uncertainty and biological variability. Being able to include uncertainty is indeed a key hallmark of simple models compared to larger models as it is easier to place such model within a statistical framework. We hope that these surprising but significant results indicate how different model strategies can be used inclusively and can help the researcher in understanding the limitations in prediction over different modeling scales.

Supplementary Material

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fphar.2015.00059/abstract

References

Beattie, K. A., Luscombe, C., Williams, G., Munoz-Muriedas, J., Gavaghan, D. J., Cui, Y., et al. (2013). Evaluation of an *in silico* cardiac safety assay: using ion channel screening data to predict QT interval changes in the rabbit ventricular wedge. *J. Pharmacol. Toxicol. Methods* 68, 88–96. doi: 10.1016/j.vascn.2013.04.004

Cavero, I., and Holzgrefe, H. (2014). Comprehensive *in vitro* Proarrhythmia Assay, a novel *in vitro/in silico* paradigm to detect ventricular proarrhythmic liability: a visionary 21st century initiative. *Expert Opin. Drug Saf.* 13, 745–758. doi: 10.1517/14740338.2014.915311

Chain, A. S. Y., Krudys, K. M., Danhof, M., and Della Pasqua, O. (2011). Assessing the probability of drug-induced QTc-interval prolongation during clinical drug development. *Clin. Pharmacol. Ther.* 90, 867–875. doi: 10.1038/clpt.2011.202

Davies, M. R., Mistry, H. B., Hussein, L., Pollard, C. E., Valentin, J.-P., Swinton, J., et al. (2012). An *in silico* canine cardiac midmyocardial action potential duration model as a tool for early drug safety assessment. *Am. J. Physiol. Heart Circ. Physiol.* 302, H1466–H1480. doi: 10.1152/ajpheart.00808.2011

Di Veroli, G. Y., Davies, M. R., Zhang, H., Abi-Gerges, N., and Boyett, M. R. (2013). High-throughput screening of drug binding dynamics to hERG improves early drug safety assessment. *Am. J. Physiol. Heart Circ. Physiol.* 304, H104–H117. doi: 10.1152/ajpheart.00511.2012

Di Veroli, G. Y., Davies, M. R., Zhang, H., Abi-Gerges, N., and Boyett, M. R. (2014). HERG inhibitors with similar potency but different binding kinetics do not pose the same proarrhythmic risk: implications for drug safety assessment. *J. Cardiovasc. Electrophysiol.* 25, 197–207. doi: 10.1111/jce.12289

Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R., and Sethna, J. P. (2007). Universally sloppy parameter sensitivities in

systems biology models. *PLoS Comput. Biol.* 3:e189. doi: 10.1371/journal.pcbi.0030189

Kramer, J., Obejero-Paz, C. A., Myatt, G., Kuryshev, Y. A., Bruening-Wright, A., Verducci, J. S., et al. (2013). MICE models: superior to the HERG model in predicting torsade de pointes. *Sci. Rep.* 3:2100. doi: 10.1038/srep02100

Lehnart, S. E., Ackerman, M. J., Benson, D. W., Brugada, R., Clancy, C. E., Donahue, J. K., et al. (2007). Inherited arrhythmias a National Heart, Lung, and Blood Institute and Office of Rare Diseases workshop consensus report about the diagnosis, phenotyping, molecular mechanisms, and therapeutic approaches for primary cardiomyopathies of gene mutations affecting ion channel function. *Circulation* 116, 2325–2345. doi: 10.1161/CIRCULATION-AHA.107.711689

Lu, H. R., Yan, G.-X., and Gallacher, D. J. (2013). A new biomarker – index of Cardiac Electrophysiological Balance (iCEB) – plays an important role in drug-induced cardiac arrhythmias: beyond QT-prolongation and Torsades de Pointes (TdPs). *J. Pharmacol. Toxicol. Methods* 68, 250–259. doi: 10.1016/j.vascn.2013.01.003

Mirams, G. R., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B. M., et al. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc. Res.* 91, 53–61. doi: 10.1093/cvr/cvr044

Mirams, G. R., Davies, M. R., Brough, S. J., Bridgland-Taylor, M. H., Cui, Y., Gavaghan, D. J., et al. (2014). Prediction of Thorough QT study results using action potential simulations based on ion channel screens. *J. Pharmacol. Toxicol. Methods* 70, 246–254. doi: 10.1016/j.vascn.2014.07.002

Pollard, C., Abi Gerges, N., Bridgland-Taylor, M., Easter, A., Hammond, T., and Valentin, J.-P. (2010). An introduction to QT interval prolongation and non-clinical approaches to assessing and reducing risk. *Br. J. Pharmacol.* 159, 12–21. doi: 10.1111/j.1476-5381.2009.00207.x

Sager, P. T., Gintant, G., Turner, J. R., Pettit, S., and Stockbridge, N. (2014). Rechanneling the cardiac proarrhythmia safety paradigm: a meeting report from the Cardiac Safety Research Consortium. *Am. Heart J.* 167, 292–300. doi: 10.1016/j.ahj.2013.11.004

Stockbridge, N., Morganroth, J., Shah, R. R., and Garnett, C. (2013). Dealing with global safety issues. *Drug Saf.* 36, 167–182. doi: 10.1007/s40264-013-0016-z

Trayanova, N. A. (2011). Whole-heart modeling applications to cardiac electrophysiology and electromechanics. *Circ. Res.* 108, 113–128. doi: 10.1161/CIRCRESAHA.110.223610

Zou, K. H., O'Malley, A. J., and Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation* 115, 654–657. doi: 10.1161/CIRCULATIONAHA.105.594929