Sex-specific classification of drug-induced Torsade de Pointes susceptibility using cardiac simulations and machine learning

Alex Fogli Iseppe[1], Haibo Ni[1], Sicheng Zhu[1], Xianwei Zhang[1], Raffaele Coppini[3], Pei-Chi Yang[2], Uma Srivatsa[4], Colleen E. Clancy[1,2], Andrew G. Edwards[1], Stefano Morotti[1], Eleonora Grandi[1]*

[1] Department of Pharmacology, University of California, Davis, CA, USA.
[2] Department of Physiology and Membrane Biology, University of California, Davis, CA, USA.
[3] Department of Neuroscience, Psychology, Drug Sciences and Child Health (NeuroFarBa), University of Florence, Italy.
[4] Department of Internal Medicine, University of California, Davis, CA, USA.

***Correspondence**

ele.grandi@gmail.com

Running Title: *Sex-specific classifiers of torsadogenic risk*

Key words: Cardiac safety pharmacology, computational modeling, machine learning classification

Abstract

Torsade de Pointes (TdP), a rare but lethal ventricular arrhythmia, is a potential cardiac side effect of drugs. To assess TdP risk, safety regulatory guidelines require to quantify the effects of new therapeutic compounds on hERG channel block *in vitro* and QT interval prolongation in vivo. Unfortunately, these have proven to be poor predictors of torsadogenic risk, and are likely to have prevented safe compounds from reaching the clinical phase. While this has stimulated numerous efforts to define new paradigms for cardiac safety, none of the recently developed strategies accounts for patient conditions. In particular, despite being a well-established independent risk factor for TdP, female sex is vastly underrepresented in both basic research and clinical studies, and thus current TdP metrics are likely biased toward the male sex. Here, we apply statistical learning to synthetic data, generated by simulating drug effects on cardiac myocyte models capturing male and female electrophysiology, to develop new sex-specific classification frameworks for TdP risk. We show that 1) TdP classifiers require different features in females vs. males; 2) male-based classifiers perform more poorly when applied to female data; 3) female-based classifier performances are largely unaffected by acute effects of hormones (i.e., during various phases of the menstrual cycle). Notably, when predicting TdP risk of intermediate drugs on female simulated data, male-biased predictive models consistently underestimate TdP risk in women. Therefore, we conclude that pipelines for preclinical cardiotoxicity risk assessment should consider sex as a key variable to avoid potentially life-threatening consequences for the female population.

1. Introduction

During drug development, promising therapeutic compounds are tested to evaluate their potential risk of inducing Torsade de Pointes (TdP), a specific form of polymorphic ventricular tachycardia that can precipitate ventricular fibrillation and cause sudden cardiac death[1]. While TdP is a very rare adverse event, amounting to less than one case out of 100,000 exposures for some non-antiarrhythmic drugs,[2] cardiac safety concerns have caused withdrawal from the market of several drugs, including antihistamines, antidepressants, chemotherapeutics, pain medications, that had been associated with TdP proclivity in patients (e.g., Cisapride and Astemizole)[2,3]. The most simple mechanistic explanation of torsadogenicity involves a reduction of the rapid delayed rectifier potassium current (I_{Kr}), carried by the human Ether-à-go-go-Related Gene (hERG) channel, which importantly contributes to

cardiac action potential (AP) repolarization[4,5]. Pharmacological block of the hERG channel, which is a very promiscuous target interacting with cardiac and non-cardiac drugs, produces AP duration (APD) and QT interval prolongation, and leads to an increased susceptibility to pro-arrhythmic events. Based on this evidence, current safety regulatory guidelines require the measurement of hERG channel block *in vitro* and QT interval prolongation *in vivo* to estimate TdP risk[6,7]. Since their adoption, these guidelines have successfully avoided that cardiotoxic drugs could endanger the welfare of people. However, it has also become apparent that these biomarkers are poor predictors of torsadogenic risk, and have in all probability prevented safe treatments from reaching the market[8,9].

In response to this problem, recent efforts have led to several proposed new paradigms for the prediction of TdP. One notable example is the Comprehensive In Vitro Proarrhythmia Assay (CiPA) initiative, an international multi-group initiative by regulatory, industry, and academic partners including the US Food and Drug Administration[10]. This paradigm relies upon the idea of combining *in vitro* studies to measure the drugs effects on each of the different types of ion channels and *in silico* models of cardiac myocyte electrophysiology to understand how these effects combine to influence cardiac function, thus creating a novel tool for TdP risk assessment of new drugs[11–13]. Mathematical models of cardiac electrophysiology, in fact, make it possible to simulate with precision extreme conditions, e.g., high drug concentrations, and to obtain insights precluded to animal experiments. Thus, computational approaches have become essential components of numerous strategies to predict torsadogenic risk[14–19]. In addition, simulated measurements extracted from the biophysical model simulations can also be fed to machine learning (ML) pipelines, as demonstrated by the Sobie group[15], with the potential to bring out mechanistic insights buried in the data that could be otherwise ignored.

To our knowledge, however, no simulation-based approach has considered any risk factor for Torsade in their predictive pipelines. An emblematic example is represented by the female sex: it is well established that women are more susceptible to Torsade than men when treated with QT-prolonging drugs,[20–22] suggesting that TdP risk classifiers could benefit from inclusion of this variable. However, female sex is highly underrepresented in both basic research[23] and clinical studies[24] involved in the drug development process, with important consequences on the identification of accurate TdP predictors. In vitro studies tend to use mostly male animals[23], raising concerns on the generalizability of findings to the whole population. This sex bias propagates onto the mathematical models of cardiac cells[25,26], which are parameterized based on male-dominated datasets. The issue of underestimating potential health risks for women is then aggravated by the fact that female sex is also underrepresented in clinical cohorts,[24] making training of classifiers harder due to the lack of reliable ground truth data.

Yang and Clancy have recapitulated *in silico* male and female ventricular human cardiac electrophysiology by incorporating experimentally determined sex- and hormone-specific differences in gene and protein expression into virtual male and female myocytes[27,28]. In this paper we combined simulations of these mathematical models with machine learning to generate sex-specific TdP risk classifiers. We simulated the effects of 59 training drugs under different conditions using *in silico* models of human ventricular myocytes with sex-specific parameterizations. We fed the resulting high-dimensional datasets of simulated biomarkers to machine learning algorithms to generate male and female classifiers of torsadogenic risk. Finally, we evaluated the effects of using sex-specific models for risk prediction on a separated set of 36 drugs, which are deemed at intermediate risk of TdP. Our results show that TdP classifiers trained on sex-specific datasets identify distinct and not interchangeable sets of optimal features, suggesting potential different drivers of drug-induced arrhythmias, and that the use of sex-biased predictive models underestimates the torsadogenic risk of drugs with intermediate risk of TdP in females, which could potentially lead to life-threatening consequences for women.

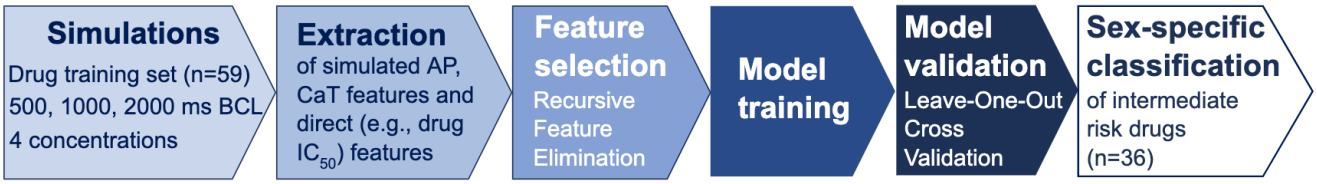

***Figure 1**: Workflow for the creation and testing of the sex-specific classifiers for torsadogenic risk.*

2. Materials and Methods

2.1 Models and simulations

We used the male and female human epicardial ventricular cardiomyocyte models developed by the Clancy lab[28], based on the O'Hara-Rudy model[29]. Some parameters of the baseline female model were modified to recapitulate observed functional sex differences in Ca^{2+}-handling[30–32]. Namely, we increased the maximal transport rate of the sodium-calcium exchanger (NCX) by 15% in the female model, and removed the originally introduced female-to-male differences in SERCA and sodium/potassium-ATPase (NKA) formulations, in agreement with experimental measurements[31,33].

To build the complete set of biomarkers used to train the ML classifiers, the virtual myocytes (with and without drug administration) were paced at a basic cycle length (BCL) of 500, 1,000, and 2,000 ms for 1,000 beats. Steady-state was confirmed as the intracellular sodium concentration had a beat-to-beat variation smaller than 0.001 mM. Each drug was simulated at multiple concentrations, ranging from 1 to 4 times their effective free therapeutic plasma concentration (EFTPC). The effects of each compound were simulated using a pore block model based on the available IC_{50} values and Hill coefficients for various ion channels (a full list is available in **Table 1**). A total of 27 biomarkers (**Table 2**) were measured on the last simulated beat for each of the 12 conditions (4 concentrations x 3 pacing frequencies). Additional biomarkers included "direct" rather than simulated measures, such as IC_{50} values for I_{Kr}, I_{Na} (fast sodium current) and I_{CaL} (L-type calcium current).

2.2 Drug dataset and labels

The dataset used in this study was obtained by combining 83 compounds analyzed in the study by the Lancaster and Sobie[15] with 12 CiPA compounds[12]. Unfortunately, we are not aware of any clinical source of torsadogenic risk categorization that takes in account the sex variable. In order to assign a unique binary label to each drug, we took advantage of the TdP risk classification available at www.crediblemeds.org[34], which reviews and analyzes adverse event reports for placing drugs in three broad categories: known, possible, and conditional risk of inducing TdP. Compounds characterized by known risk of inducing TdP in CredibleMeds were considered TdP^+ in our analysis, while safe compounds (i.e., not included in any of the CredibleMeds categories) were considered TdP^-. Being sex one of the factors causing the inclusion of drugs in the possible or conditional risk categories, it would have been inaccurate to assign these compounds to a specific class in our binary classification problem. Therefore, we did not use these drugs in the training phase of our classifiers. A similar selection was performed on the CiPA dataset, which separates the compounds in high, intermediate, and low torsadogenic risk. Drugs classified differently by CredibleMeds and CiPA were excluded from the training dataset. Out of the 95 drugs in our initial dataset, 59 drugs met all the requirements. The remaining 36 drugs, which are more likely to be associated with sex-specific effects, were then used to test our TdP classifiers and investigate the consequences of performing male-sex-biased vs. sex-specific predictions.

2.3 Machine Learning

A schematic diagram describing the complete workflow used in the generation of the ML classifiers for TdP risk is shown in **Fig 1**. First, the simulated recordings are visually inspected for the presence of early after depolarizations (EADs) or alternans in the last 3 beats. If proarrhythmic events are observed, e.g., EADs or repolarization failure that would alter the values of APD, the biomarkers for that specific drug-

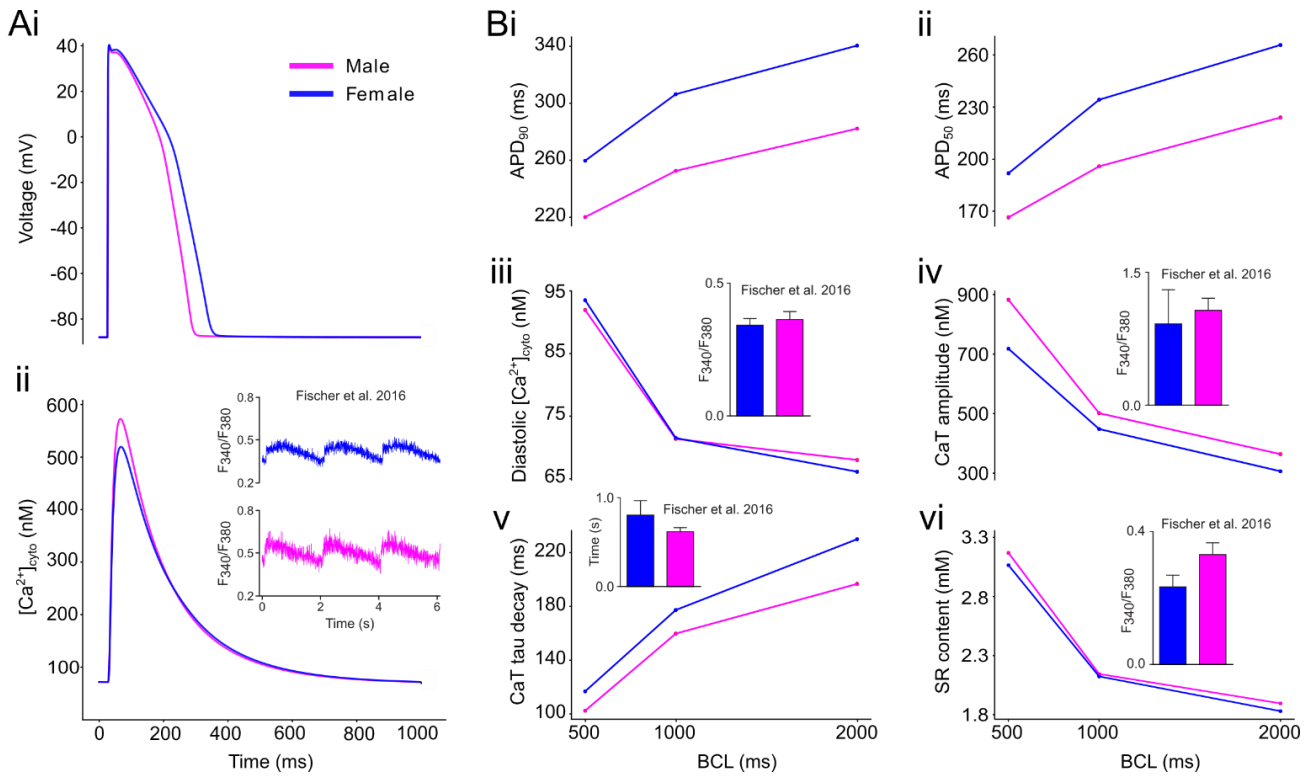

Figure 2: Sex-specific biophysical models of cardiac electrophysiology and Ca^{2+} handling. ***A****: simulated AP* ***(i)*** *and Ca^{2+} transient* ***(ii)*** *traces pacing at a BCL of 1,000 ms the baseline models of male (magenta) and female (blue) ventricular epicardial cardiomyocyte.* ***B****: Rate dependency of simulated biomarkers: measurements of voltage-* ***(i,ii)*** *and Ca^{2+}-related* ***(iii-vi)*** *biomarkers as a function of the BCL tested. Insets show representative and summary data for Ca^{2+} transient characteristics measured in myocytes from patients with cardiac hypertrophy (reproduced with permission from Fischer et al. 2016)*[41].

pacing frequency combination are imputed as following. The average value of each biomarker in the TdP^+ and TdP^- groups is measured. The drug will receive the (maximum) minimum value of its TdP outcome group if the average value of its group is (larger) smaller than the value for the other group. For example, TdP^+ drugs have a larger average value for APD_{90} than TdP^- drugs. As a consequence, the APD_{90} value of a TdP^+ drug producing EADs in a specific condition (e.g., Ibutilide at 4x EFTPC and 500 ms of BCL) will be the maximum APD_{90} measured in that condition among all the TdP^+ drugs that do not produce EADs. To offset the differences of the sex-specific baseline models, the measured biomarkers in drug-free conditions are subtracted to those in response to drug administration. Lastly, since biomarkers have different scales, the data are standardized for better ML performances and results interpretability.

In order to select the biomarkers contributing to the most accurate prediction of torsadogenic outcome, we adopted a recursive feature elimination (RFE) algorithm. At each step of the algorithm, a new classifier is trained with the available features and its regularization parameter are tuned to achieve the best performances. The predictive power of the model is quantified in the RFE algorithm using the Matthew's Correlation Coefficient (MCC) measured through Leave-One-Out Cross-Validation (LOO-CV). The features are then ranked based on their importance for the classification task, and the least important feature is discarded by the training dataset. The process is repeated until all the features are eliminated. The best candidate is the ML classifier characterized by the highest MCC using the smallest set of features. We calculated the area under the receiver operating characteristic Curve (AUC) and the F1 score as additional performance metrics. For the sake of feature interpretability, we limited the ML modeling algorithms to logistic regression and support vector machine (SVM) with linear kernel. All the results shown

here were obtained using SVM models, which outperformed logistic regression models on our datasets.

2.4 Simulation and data analysis software, numerical method, and code availability

The Yang and Clancy model[28] code was implemented in C++. We utilized the male model as is and modified the female model as described above. The ODEs were solved using a combination of Forward Euler and Rush-Larsen scheme[35], as done in the original O'Hara-Rudy model[29] and implemented with a variable time step (dt = 0.025 or 0.005 ms). The data processing, RFE algorithm, and machine learning modeling were implemented in Python using the packages Numpy[36], Pandas[37], Scikit-Learn[38], and Hyperopt[39].

All simulations and data analyses were performed on a desktop server: HP Z2 Tower G4, Intel(R) Core(TM) i7-8700K @ 3.20GHz 6CPUs (12 threads) + 16GB; and a computing cluster with Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz 28 CPUs (56 threads) + 132GB.

Source code and documentation are freely available at http://elegrandi.wixsite.com/grandilab/downloads and https://github.com/drgrandilab.

3. Results

3.1 Sex-specific biophysical models for action potential and Ca^{2+} transient

The Yang and Clancy model[28] well recapitulates the prolonged AP observed in females vs. males[40]. Conversely, we noted that the predicted larger Ca^{2+} loading and transient amplitude in the female model does not match the sex differences measured in rodents[30,32]. Our updated female model, with modifications in Ca^{2+}-handling processes described in the Methods, captures these sex-specific differences (**Fig 2Aii**). Namely, the female Ca^{2+} transient amplitude is modestly reduced (**Fig 2Biv**) and decays slightly more slowly (**Fig 2Bv**) than in male. At the same time, the diastolic Ca^{2+} concentration and SR content do not appreciably differ in male and female (**Fig 2Aiii**,**Avi**). These results are similar to Ca^{2+} measurements in myocytes from patients with cardiac hypertrophy (**Fig 2 *insets***), whereby no significant sex differences were detected in systolic or diastolic Ca^{2+} levels, decay rate, and SR Ca^{2+}content[41]. Notably, AP properties are very modestly affected by the modifications we introduced in the model parameters, thus preserving the typical differences in repolarization between the two sexes in the original Yang and Clancy model[28] (**Fig 2Ai**,**Bi**,**Bii**).

3.2 Sex-specific TdP risk classifiers

With the original male and updated female models, we simulated the effects of the 59 training drugs under different pacing and drug regimen conditions (see *Methods*). Representative simulated traces reflecting the ranges of variations induced by the drugs on the AP and Ca^{2+} transient are shown in **Fig 3A** (male) and **B** (female). For each simulated condition, several biomarkers (**Table 2**) were extracted and, together with the IC_{50} values available for each drug, generated the final feature datasets used by the RFE algorithm. This process led to optimized male- and female-specific TdP classifiers.

The final sets of features and relative weights selected and used by the best performing male and female TdP risk classifiers are illustrated in **Fig 3C** and **3E**, respectively. TdP risk predictions for the male dataset are based on the current integrals of the inward rectifier I_{K1} and late Na^+ current I_{NaL}, and the diastolic $[Ca^{2+}]$. With LOO-CV, this ML model can correctly classify 54 out of 59 drugs, whereby 4 proarrhythmic drugs are predicted safe (4 false negatives, FNs – Amiodarone 1, Amiodarone 2, Cilostazol, Donepezil, where Amiodarone is simulated using different available IC_{50} as shown in **Table 1**) and 1 safe drug is predicted harmful (1 false positive, FP – Prenylamine). The best performing female classifier is built on a set of 5 features. Four of the features are measurements extracted from the Ca^{2+} transient (decay times, integral, and diastolic $[Ca^{2+}]$), while the fifth one is the value of IC_{50} for the hERG channel. The model optimized on the female dataset misclassifies 3 drugs: 1 FN (Procainamide) and 2 FPs (Ajmaline, Prenylamine). Despite the different set of features involved, both classifiers clearly demonstrate good ability in distinguishing torsadogenic from safe drugs.

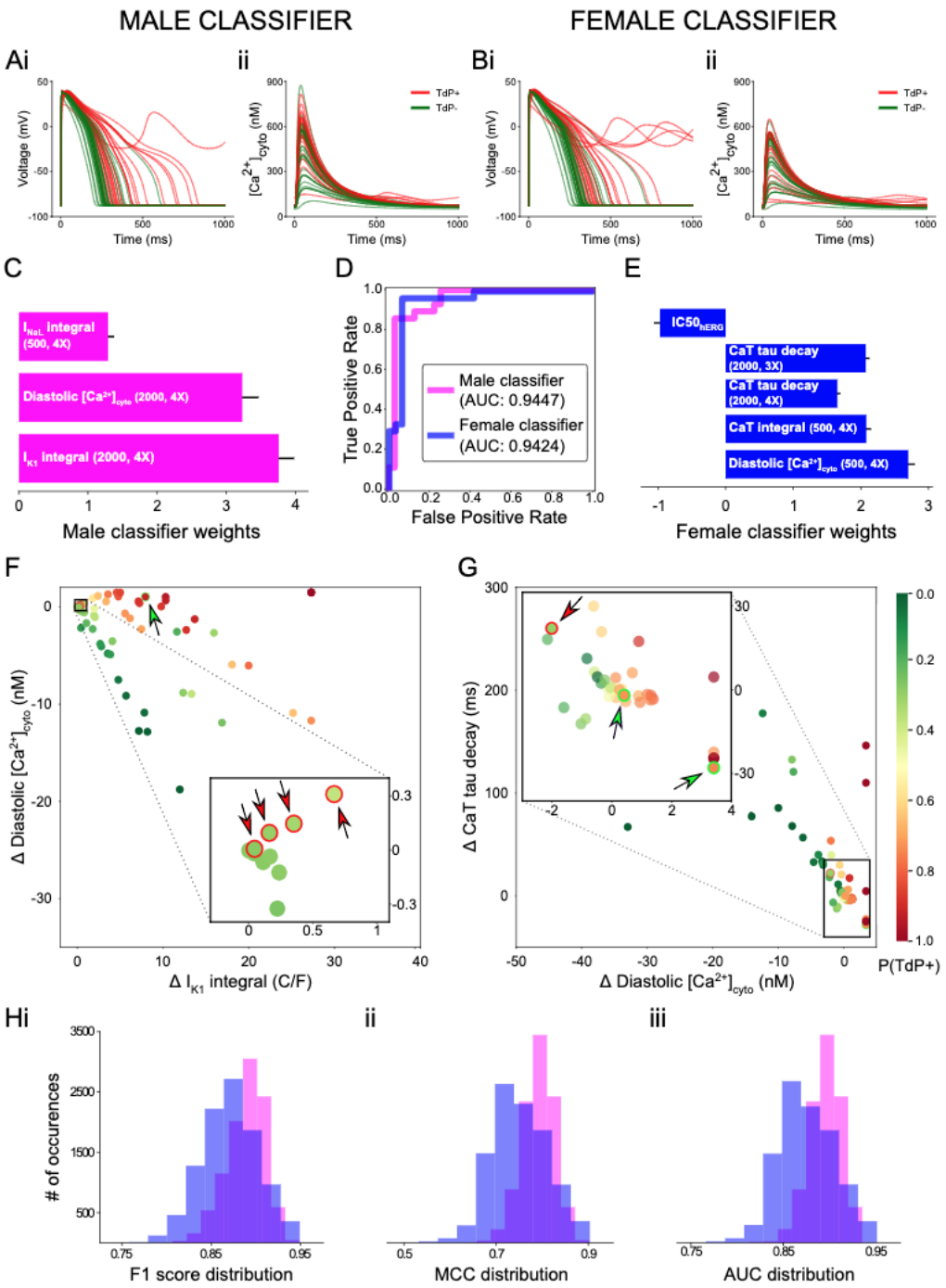

***Figure 3**: **Sex-specific TdP risk classifiers**. **A,B**: simulated effects of the 59 training drugs on the AP (i) and Ca^{2+} transient (ii) with male (**A**) and female (**B**) biophysical myocyte models paced at a BCL of 1,000 ms. Traces belonging to drugs TdP^{+} and TdP^{-} are colored in red and green, respectively. **C,E**: best performing set of features selected through RFE and used to build the male (**C**) and female (**E**) TdP classifiers. Uncertainty of the feature weights is measured using LOO-CV (mean + SD). **D**: Receiver operating characteristic curve for male (magenta) and female (blue) TdP classifiers. Area under the curve (AUC) is 0.9447 and 0.9424 for male and female, respectively. **F,G**: scatterplot of the training drugs created using the two features with the largest weight for the male (**F**) and female (**G**) TdP classifiers. The estimated TdP risk probability for each drug is indicated by the color of the filling. The misclassified drugs are indicated by arrows, and the color of the arrow and the stroke specifies the right class (green for TdP^{-}, red for TdP^{+}). **H**: distributions of performance metrics for the male (magenta) and female (blue) TdP classifiers after injecting random normally-distributed (μ=0, σ=0.1) noise to the original data (10,000 repetitions).*

A

	Female classifier	Male classifier	Male classifier retrained
F1 score	0.9492	0.7693	0.898
MCC	0.8987	0.5945	0.7972
AUC	0.9424	0.9159	0.9217
FN	1	12	4
FP	2	1	2
Misclassified drugs	Ajmaline, Prenylamine, Procainamide	Amiodarone 1, Amiodarone 2, Bepridil 2, Bepridil CiPA, Cilostazol, Donepezil, Flecainide, Halofantrine, Prenylamine, Procainamide, Quinidine 1, Sotalol, Thioridazine 1	Ajmaline, Amiodarone 1, Amiodarone 2, Cilostazol, Donepezil, Prenylamine

B

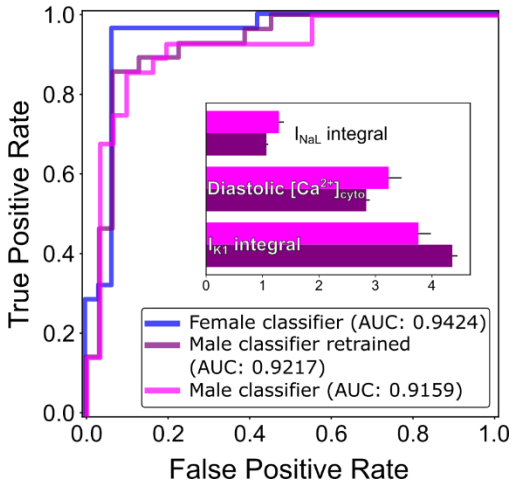

Figure 4*: *Predictions of TdP risk on female data with male-specific features*. *A*: *TdP classifier performances are evaluated on the dataset created with the female biophysical model. The results obtained with the female classifier (left column) are compared with the ones of the original male classifier (center column) and of a male classifier retrained on the female dataset (right column). B: Receiver operating characteristic curve for female (blue), original male (blue), and retrained male (purple) TdP classifiers applied on female data; area under the curve (AUC) is 0.9424, 0.9159, and 0.9217, respectively. In the inset, bar plot comparing the weights of the original (magenta) and retrained (purple) male classifiers. Uncertainty of the feature weights is measured using LOO-CV (mean + SD).

To evaluate model performances with commonly used metrics, we calculated F1 scores of 0.9148 and 0.9492, and MCCs of 0.8336 and 0.8987 for male and female TdP risk classifier, respectively (**Table 3**). Robust performances were confirmed by AUC values of 0.9447 and 0.9424 (**Fig 3D**) values for male and female classifier, respectively.

To visualize the TdP risk predictions of the classifiers, we created 2D scatterplots of the training compounds using the two most influential features used by each model (**Fig 3F** for male, **Fig 3G** for female). In this representation, the farther the features from the drug-free condition at coordinates (0,0), i.e., the larger the drug effect on the features, the more extreme are the predicted probabilities (darker colors). Interestingly, most of the misclassified drugs are located in the features space around the drug-free condition (insets of **Fig 3F** and **3G**), and are associated with probabilities close to the classification threshold of 0.5. To evaluate the robustness of our classifiers, we added random normally-distributed noise (μ=0, σ=0.1) to the features and evaluated their performances by repeating this procedure 10,000 times (**Fig 3H**). The noise injection tends to increase the misclassification rate of the classifiers, whereby the drugs that are more frequently misclassified are those located in close proximity to the decision boundary of the SVM (r = -0.7653 and -0.8808 for male and female, respectively). However, despite being negatively affected, the performances of the sex-specific classifiers remain satisfactory even in presence of confounding noise in the data.

3.3 Testing the male classifier on female data

To verify our contention that the creation of sex-specific classifiers is indeed critical to obtain accurate predictions, we evaluated the performance of the male classifier in predicting TdP risk of the training drugs applied to the female simulated features. When applied to female data, the predictive accuracy of the male classifier dropped, producing 1 FP and 12 FNs (**Fig 4A**). One possible explanation for the poor performances could be that the male classifier was trained on a different dataset. To test this idea, we retrained the male classifier using the data generated from the simulations with the female biophysical model (weights are compared in the insert of **Fig 4B**)

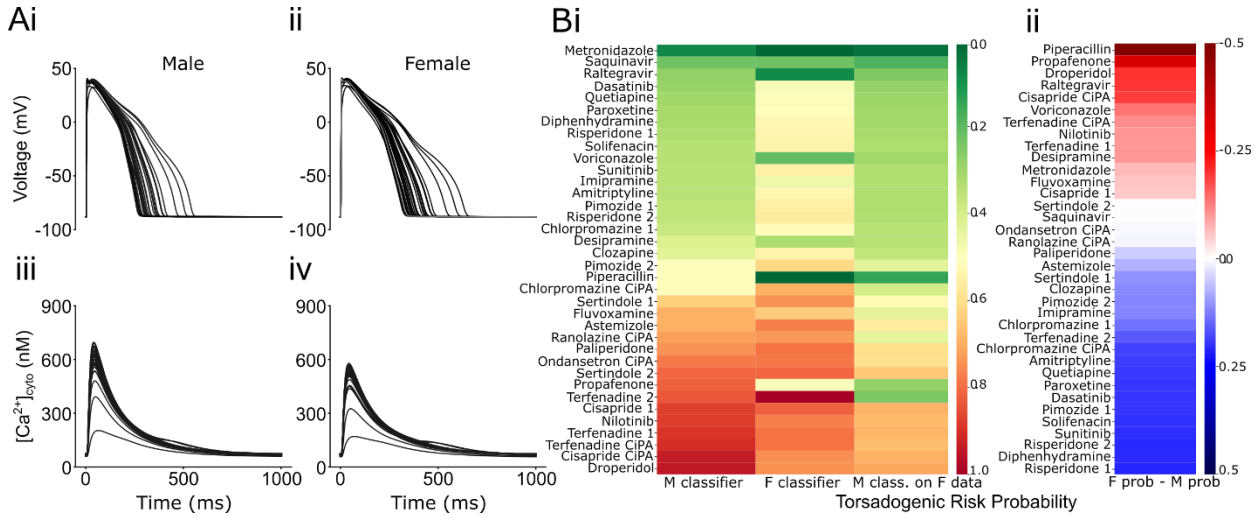

***Figure 5**: **Sex-specific TdP predictions for intermediate-risk drugs**. **A**: simulated effects of the 36 intermediate drugs on the AP (**i,ii**) and Ca^{2+} transient (iii,iv) on the male (left) and female (right) biophysical models paced at a BCL of 1,000 ms. **B**: heatmaps of the predicted torsadogenic risk probability estimated by the sex specific TdP classifiers. (**i**) Left and center columns show the predictions obtained in male and female, respectively. The torsadogenic risk probabilities predicted by the male TdP classifier on the female dataset are visualized in the right column. Drugs are sorted by the risk probability predicted by the male TdP classifier. (**ii**) Difference between torsadogenic risk probabilities predicted by female and male TdP classifiers.*

and re-evaluated the performances. The retrained male ML model still produced more misclassified drugs compared to our female classifier (**Fig 4A**). This result demonstrates that the set of features used by the female classifier has superior predictive ability for drug-induced arrhythmias in females, and confirms the need of considering sex when estimating drug-induced torsadogenic risk in preclinical compounds (and in patients).

3.4 Sex-specific prediction of drugs with intermediate risk of TdP

From the initial list of drugs in our possession, 36 drugs had not been included in the training phase for belonging to the so-called intermediate torsadogenic risk category. In these drugs, TdP risk is associated with the presence of one or more risk factors. Indeed, when simulating these compounds with the sex-specific cardiomyocyte models, the observed changes on the AP (**Fig 5Ai,ii**) and Ca^{2+} transient (**Fig 5Aiii,iv**) are milder compared to the drugs with higher risk (**Fig 3A,B**). In order to explore how sex could affect the estimated torsadogenic risk of the drugs in the intermediate category, we applied the sex-specific TdP classifiers to the simulated male and female features and compared the predicted TdP probabilities (**Fig 5Bi**). In general, intermediate TdP risk drugs tend to have more dangerous outcomes in women, whereby a larger number of compounds is predicted to have higher probability of TdP in women (**Fig Bii**). Notably, if the male classifier is used on female data (right columns, **Fig 5Bi**), the predicted risk of the compounds is consistently underestimated, confirming the results obtained with the training dataset.

3.5 Effect of hormones on TdP risk prediction

It is well-established in the literature that circulating levels of hormones affect cardiac electrophysiology[42]. To test how the performances of the female TdP classifier are influenced by hormones, we simulated the drug effects during the different phases of the menstrual cycle. Notably, the performances of the female classifier are almost unaltered by the hormonal effects (**Fig 6A**), with some modest changes in prediction in the late follicular phase. This is explained by hormones having minimal consequences on the Ca^{2+} transient, which strongly influences the female

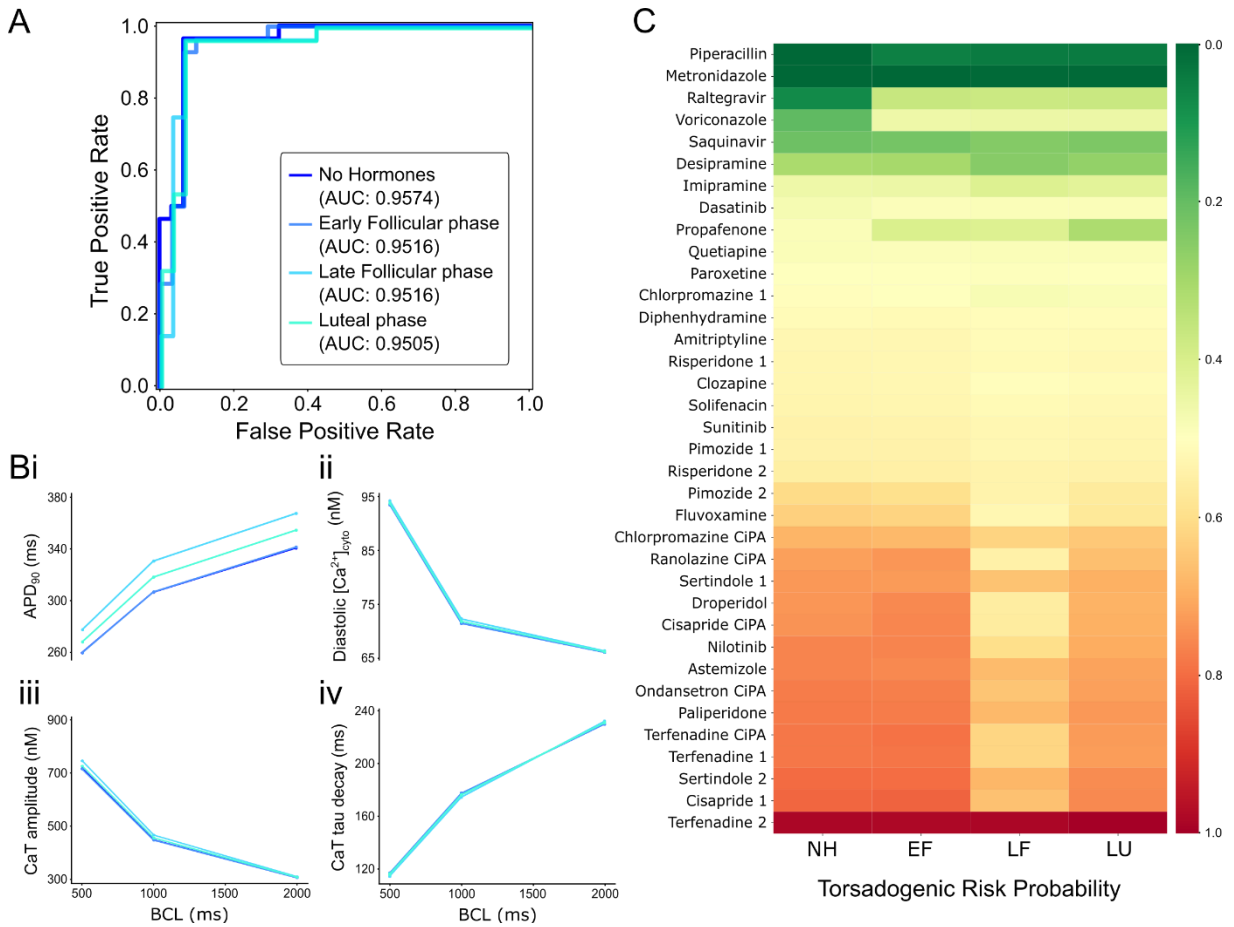

***Figure 6**: **Effects of sex hormone fluctuations on female TdP classifier predictions**. **A**: Receiver operating characteristic curve for female TdP classifier applied on training drugs dataset in absence of hormonal effects (navy blue) or simulating early follicular (EF, azure blue), late follicular (LF, sky blue), and luteal (LU, turquoise) phases. Area under the curve (AUC) is 0.9574, 0.9516, 0.9516, and 0.9505, respectively. **B**: Rate dependency of biomarkers simulated during different menstrual phases: measurements of voltage- **(i)** and Ca^{2+}-related **(ii-iv)** biomarkers as a function of the BCL tested. **C**: heatmap of the predicted torsadogenic risk probabilities of intermediate TdP risk drugs simulated during different menstrual phases.*

predictions, while prolonging the AP duration considerably (**Fig 6B**). The consistency of the female classifier predictions in response to hormonal perturbations is also reflected in the probabilities forecasted for the drugs at intermediate TdP risk (**Fig 6C**). Similar results were obtained for the male classifier simulating the effects of testosterone (not shown). Taken together, these results show that this torsadogenic classifier is robust to acute changes in the levels of the sex hormones.

4. Discussion

Despite its important role in susceptibility to Torsade de Pointes[20–22], sex is rarely considered when developing predictive frameworks for torsadogenic risk. In the current study, we updated a published model of ventricular cardiomyocyte with sex-specific parameterizations[28] to include experimentally observed differences in Ca^{2+} handling[30–32], and used it to simulate the effects of drugs belonging to different TdP risk categories. We then fed the simulated data to machine learning algorithms and generated sex-specific TdP risk classifiers. We showed that: (i) classifiers trained on data reflecting male and female electrophysiological properties are built on distinct and not interchangeable set of features, (ii) male classifiers underestimate the torsadogenic risk in females, (iii) female classifier predictions are robust to changes in sex hormone fluctuations during the menstrual cycle. Taken

together, these results confirm the need of including sex when estimating the torsadogenic risk of a drug and provide a new tool to aid in this investigation.

4.1 Prediction of TdP risk

Our work is preceded by a number of important studies that utilized *in silico* approaches to identify biomarkers predictive for Torsade. In 2011, Mirams et al.[14] used cardiac models to investigate the role of multi-channel effects on forecasting TdP risk categories. They showed that simulating the effects of 31 drugs on I_{Na} , I_{CaL} and I_{Kr} allowed prediction of TdP risk better than that based on I_{Kr} block only. Similar conclusions were reached in the study by Kramer et al.[43], who trained logistic regression models on combinations of pharmacodynamic parameters. The necessity of considering how drugs affect multiple targets is also one of the founding pillars of the CiPA initiative[10]. Since its creation, this multi-group effort has built various tools for improving the simulations of drug effects[11,12,44] and developed the *Torsade Metric Score*, an original metric which, in its most recent update[13], uses the sum of the integrals of 4 different ionic currents to order 28 compounds by their TdP risk. Our approach, based on a combination of biophysically-detailed simulations and machine learning, is inspired by the seminal work of Lancaster and Sobie[15]. The study demonstrated that the performance of a SVM classifier trained on drug effects on APD_{50} and diastolic Ca^{2+} levels was comparable to that built on an entire set of biomarkers, and outperformed existing TdP risk predictive models. The specificity of this predictive model was later improved by Krogh-Madsen et al.[16] with the adoption of a cell model optimized on both clinical APD data and intracellular ionic concentrations. The Clancy group extensively demonstrated the use of Triangulation, Reverse-use dependence, Instability and Dispersion (i.e. TRIaD)[45,46], proposed by Hondeghem et al.[47], a combined metric suggesting that the kinetics of repolarization and the instability of AP prolongation are more critical than AP lengthening itself. Other groups have suggested that more direct metrics can still contain enough information for satisfactory TdP risk predictions[48]. The *TdP score* developed by the Rodriguez group, for example, relies only on the count of repolarization abnormalities observed during drug applications in experimentally calibrated populations of cardiomyocyte models[18]. This estimator has indeed demonstrated promising performances, which have been further increased by adding simulation-derived measurements to the score calculation[18]. In our study, we used both direct features (i.e., IC_{50} values for I_{Kr}, I_{CaL}, and I_{Na}) and simulated biomarkers as inputs to the ML models. For the first time, we utilized sex-specific ionic models, based on the seminal work of the Clancy lab[27,28,49] (accounting for both chronic and acute effects of sex hormones on cardiac electrophysiology) for creating sex-specific TdP risk classifiers.

4.2 Sex-specific model of ventricular electrophysiology and Ca^{2+} handling

The female Yang and Clancy[28] model was parameterized using differences in expression levels of ion channels and transporters measured in cardiac tissue explanted from female vs. male patients[50]. In addition, the authors implemented the effects of sexual hormones on three ionic currents (I_{Kr}, I_{Ks}, I_{CaL}). While the resulting parameterizations accurately recapitulate the well documented differences in AP (and QT) properties in the basic research and clinical literature[40,51], less is known about sex differences in Ca^{2+} handling in humans. Experimental studies in rodents revealed smaller CaT amplitudes in female animals, associated with a decreased excitation-contraction coupling gain[31,32]. Interestingly, L-type Ca^{2+} current, diastolic [Ca^{2+}] and SR Ca^{2+} content are not significantly different between sexes[31,32]. The simulated Ca^{2+} levels of the original female model were unusually high, particularly at faster frequencies and compared to the simulated male Ca^{2+} levels. Therefore, we modified some parameters affecting the Ca^{2+}-handling processes in the female model (see Methods) and obtained CaTs in line with the differences reported in literature. Intriguingly, the outputs of our updated model more closely resemble the sex-specific Ca^{2+} recordings obtained by Fischer et al.[41] from isolated ventricular myocyte of hypertrophic patients (**Fig 2**). Information on the

sex-specific differences in Ca^{2+} handling from non-diseased patients would be highly desirable, but we are not aware of any available data.

4.3 Sex-specific TdP biomarkers

Here we propose for the first time the adoption of sex-specific torsadogenic risk predictors. Through an RFE algorithm, we obtained male and female classifiers built on distinct minimal set of best performing features. Importantly, the RFE iterative process does not require any manual intervention of the user, leading to unbiased results. From the analysis of the most predictive biomarkers (**Fig 3C**), we found that the male classifier resembles the SVM classifier developed by Lancaster and Sobie[15]. In both cases, TdP classification is based on information about the changes in (1) myocyte diastolic Ca^{2+} concentration and (2) one or more metrics related to the degree of drug-induced AP prolongation. Diastolic Ca^{2+} levels are an index of cell Ca^{2+} loading. In fact, we found that this biomarker is highly correlated with SR Ca^{2+} content and Ca^{2+} transient amplitude (r = 0.9803 and 0.8476, respectively). The integral of I_{K1} used by our male classifier is highly correlated with APD_{90}/AP triangulation (r = 0.8872 and 0.8946 respectively). Accordingly, the scatterplot shown in **Fig 3F** looks similar to the one published by the Sobie group (see Fig 3b in Lancaster and Sobie, 2016[15]).

The female classifier, on the contrary, selected hERG IC_{50} and several features related to Ca^{2+}-handling processes, whereby diastolic Ca^{2+} concentration, integral of CaT, and CaT decay time constant are all positively associated with torsadogenic risk (**Fig 3E**). Altered Ca^{2+} homeostasis has been shown to be an important determinant of Torsade susceptibility. An elevated Ca^{2+} influx through L-type Ca^{2+} channels (e.g., mediated by AP prolongation), for example, can lead to a stronger activation of NCX, which in turn may reactivate the I_{CaL} and promote the generation of EADs[52]. The enhancement of NCX activity could be exacerbated even more in women, where the expression levels are higher in physiological condition[33]. While a detailed mechanistic interpretation of the necessity of different CaT-related features for TdP prediction in female is still undetermined, a deeper exploration through ad hoc experiments is advisable.

4.3 Ground truth classification

We are not aware of any clinical drug classification accounting for sex-specific TdP risk. To solve this critical lack of ground truth, we operated under the assumption that drugs generally considered safe (i.e., missing in the CredibleMeds[34] database, which contains a continuously updated database of therapeutic compounds categorized by the strength of the association between their use and a torsadogenic outcome) are so in both sexes. Conversely, drugs clearly associated with Torsade (i.e., inserted in the "Know Risk of TdP" category), are more likely to be so in both sexes. On the contrary, drugs associated with an intermediate risk of causing TdP (i.e., inserted in the "Possible" or "Conditional Risk of TdP" categories) are more likely to change risk categorization depending on sex. Therefore, the latter drugs were therefore not used to build our classifiers.

Based on the ground truth adopted here, our sex-specific TdP classifiers have shown remarkable predictive powers (**Fig 3D,H**, **Table 3**). In addition, analyzing in more detail the misclassified drugs, we note that both male and female ML models produced an erroneous classification of the drug Prenylamine. This compound, missing in the CredibleMeds[34] and CiPA[10] databases, has been labeled as TdP^- in our pipeline. However, prenylamine has been reported to cause Torsade in patients[53], hence it is labeled as torsadogenic by categorization systems different from the ones we used[54]. Notably, the classifiers we developed were able to detect the torsadogenicity of the compound despite the wrong label. A similar conclusion can be reached exploring the clinical literature for ajmaline[55].

4.4 Predictions on intermediate drugs

Our classification results with the intermediate risk drugs identified multiple therapeutic compounds that are predicted safe in men and torsadogenic in women (**Fig 5B**), which is in general agreement with the increased susceptibility to TdP in females. As drug-induced TdP is a rare event, prospective studies to evaluate the TdP

risk factors are difficult to design and would require very large patient cohorts. We analyzed retrospective reviews of case reports, and notably, found a nice correspondence in the predictions for some specific compounds with observations published in the clinical literature. Indeed, female sex is a risk factor for documented TdP episodes associated with the use of Risperidone[56,57], Sunitinib[58], Paroxetine[59], and Quetiapine[60], which are among the intermediate-risk therapeutic compounds with the largest selectivity for women based on our estimated torsadogenic risk.

It is also important to understand sex-differences in TdP outcome in high-risk drugs, which also demonstrate female sex-prevalence, as demonstrated for quinidine, amiodarone, sotalol, disopyramide, bepridil, prenylamine (but not procainamide)[20].

4.5 Limitations

It is important to recognize that even at comparable drug dosages, drug exposure may vary between women and men owing to differences in absorption, distribution, metabolism and excretion that could explain higher TdP risk in women. The increased risk is not fully explained by sex difference in drug plasma levels[61], though cellular concentrations of a same systemic drug dose can vary across individuals and between sexes. Thus, future studies should account for both (population) pharmacokinetic and pharmacodynamic drug interactions. Experimental differences in pharmacokinetics observed between men and women have frequently been attributed to bodyweight differences and thus might be addressed by appropriate adjustment of dosage by body weight. In fact, in a cohort of more than 200 patients a statistically significant prevalence of dofetilide dose reduction or discontinuation was found in female vs. male patients mostly due to QTc prolongation, although no TdP cases were reported[62]. On the other hand, the occurrence of TdP was not associated to any critical serum drug level of quinidine[63].

Here, we showed that while "chronic" sex differences affect the predicted susceptibility to TdP[42], varying levels of sex hormones, e.g., during the menstrual cycle, do not impair the performance of our TdP classifiers (**Fig 6**). Several studies have reported the influences of circadian rhythms on cardiac physiology[64,65]. Circadian variation in the levels of hormones, heart rate, and drug metabolism could undoubtedly interact and influence the response of the heart to drugs, as reported for the differential QT prolongation caused by Risperidone[66] or Moxifloxacin[67] during the 24-hour period.

5. Conclusions

We have extracted sex-specific features to classify the torsadogenic risk of drugs using simulated data resembling male and female drug responses. We demonstrated that the predictive features are not interchangeable between the sexes, and are robust to noise and to variations in acute hormone effects. Taken all together, our results indicate the need of considering sex when developing and applying torsadogenic classifiers, to obtain more accurate predictions and provide safer therapeutic treatments to patients.

Clinical risk assessment and trials suggest that besides sex, other patient conditions, i.e., age, disease, electrolyte imbalance[68], interaction with other drugs[69] should all be taken into account in the evaluation TdP. Data gathered from experiment in human induced pluripotent stem cells-derived cardiomyocytes obtained from male and female cell lines have been recently proven useful for investigating the sex-differences in torsadogenicity[70], suggesting that these patient-derived cells could be used to guide new models and paradigms for safety pharmacology[19] accounting for patient conditions.

6. Funding

This work was supported by the National Institutes of Health (NIH) Stimulating Peripheral Activity to Relieve Conditions grant 1OT2OD026580 (EG and CC); the National Heart, Lung, and Blood Institute grants R01HL131517, R01HL141214, P01HL141084 (EG), and R00HL138160 (SM); the American Heart Association Postdoctoral Fellowship 20POST35120462 (HN), Predoctoral Fellowship 20PRE35120465 (XZ), and grant 15SDG24910015 (EG); the UC Davis School of Medicine Dean's Fellow award (EG) and Academic Senate grant (EG and US); the Health and Environmental Sciences Institute grant U01 FD006676-01 (AE).

References

1. Dessertenne, F. La tachycardie ventriculaire a deux foyers opposes variables. *Arch. Mal. Coeur Vaiss.* **59**, 263–272 (1966).
2. Haverkamp, W. *et al.* The potential for QT prolongation and proarrhythmia by non-antiarrhythmic drugs: clinical and regulatory implications. Report on a Policy Conference of the European Society of Cardiology. *Eur. Heart J.* **21**, 1216–1231 (2000).
3. Food and Drug Administration. Additions and Modifications to the List of Drug Products That Have Been Withdrawn or Removed From the Market for Reasons of Safety or Effectiveness. Final rule. *Fed. Regist.* **79**, 37687–37696 (2014).
4. Sanguinetti, M. C., Jiang, C., Curran, M. E. & Keating, M. T. A mechanistic link between an inherited and an acquird cardiac arrthytmia: HERG encodes the IKr potassium channel. *Cell* **81**, 299–307 (1995).
5. Tristani-Firouzi, M., Chen, J., Mitcheson, J. S. & Sanguinetti, M. C. Molecular biology of K+ channels and their role in cardiac arrhythmias11Am J Med. 2001;110-50-59. *Am. J. Med.* **110**, 50–59 (2001).
6. Food and Drug Administration. International Conference on Harmonisation; Guidance on S7B Nonclinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals; Availability. *Fed. Regist.* **70**, 61133–4 (2005).
7. Food and Drug Administration. International Conference on Harmonisation; guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs; availability. Notice. *Fed. Regist.* **70**, 61134–5 (2005).
8. Sager, P. T. Key clinical considerations for demonstrating the utility of preclinical models to predict clinical drug-induced torsades de pointes. *Br. J. Pharmacol.* **154**, 1544–1549 (2008).
9. Gintant, G. An evaluation of hERG current assay performance: Translating preclinical safety studies to clinical QT prolongation. *Pharmacol. Ther.* **129**, 109–119 (2011).
10. Sager, P. T., Gintant, G., Turner, J. R., Pettit, S. & Stockbridge, N. Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the Cardiac Safety Research Consortium. *Am. Heart J.* **167**, 292–300 (2014).
11. Dutta, S. *et al.* Optimization of an In silico Cardiac Cell Model for Proarrhythmia Risk Assessment. *Front. Physiol.* **8**, 1–15 (2017).
12. 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**, e004628 (2017).
13. Li, Z. *et al.* Assessment of an In Silico Mechanistic Model for Proarrhythmia Risk Prediction Under the Ci PA Initiative. *Clin. Pharmacol. Ther.* **105**, 466–475 (2019).
14. Mirams, G. R. *et al.* Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc. Res.* **91**, 53–61 (2011).
15. Lancaster, M. C. & Sobie, E. Improved Prediction of Drug-Induced Torsades de Pointes Through Simulations of Dynamics and Machine Learning Algorithms. *Clin. Pharmacol. Ther.* **100**, 371–379 (2016).
16. Krogh-Madsen, T., Jacobson, A. F., Ortega, F. A. & Christini, D. J. Global Optimization of Ventricular Myocyte Model to Multi-Variable Objective Improves Predictions of Drug-Induced Torsades de Pointes. *Front. Physiol.* **8**, 1–10 (2017).
17. Passini, E. *et al.* Human In Silico Drug Trials Demonstrate Higher Accuracy than

Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. *Front. Physiol.* **8**, 1–15 (2017).
18. Passini, E. *et al.* Drug-induced shortening of the electromechanical window is an effective biomarker for in silico prediction of clinical risk of arrhythmias. *Br. J. Pharmacol.* bph.14786 (2019). doi:10.1111/bph.14786
19. Li, Z. *et al.* General Principles for the Validation of Proarrhythmia Risk Prediction Models: An Extension of the CiPA In Silico Strategy. *Clin. Pharmacol. Ther.* **107**, 102–111 (2020).
20. Makkar, R. R., Fromm, B. S., Russell, T. S., Meissner, M. D. & Lehmann, M. H. Female Gender as a Risk Factor for Torsades de Pointes Associated With Cardiovascular Drugs. *JAMA J. Am. Med. Assoc.* **270**, 2590 (1993).
21. Bednar, M. M., Harrigan, E. P. & Ruskin, J. N. Torsades de pointes associated with nonantiarrhythmic drugs and observations on gender and QTc. *Am. J. Cardiol.* **89**, 1316–1319 (2002).
22. Chorin, E. *et al.* Female gender as independent risk factor of torsades de pointes during acquired atrioventricular block. *Hear. Rhythm* **14**, 90–95 (2017).
23. Flórez-Vargas, O. *et al.* Bias in the reporting of sex and age in biomedical research on mouse models. *Elife* **5**, 1–14 (2016).
24. Vitale, C. *et al.* Under-representation of elderly and women in clinical trials. *Int. J. Cardiol.* **232**, 216–221 (2017).
25. Zucker, I. & Beery, A. K. Males still dominate animal studies. *Nature* **465**, 690–690 (2010).
26. Ramirez, F. D. *et al.* Sex Bias Is Increasingly Prevalent in Preclinical Cardiovascular Research: Implications for Translational Medicine and Health Equity for Women. *Circulation* **135**, 625–626 (2017).
27. Yang, P.-C., Kurokawa, J., Furukawa, T. & Clancy, C. E. Acute Effects of Sex Steroid Hormones on Susceptibility to Cardiac Arrhythmias: A Simulation Study. *PLoS Comput. Biol.* **6**, e1000658 (2010).
28. Yang, P.-C. & Clancy, C. E. In silico Prediction of Sex-Based Differences in Human Susceptibility to Cardiac Ventricular Tachyarrhythmias. *Front. Physiol.* **3**, 1–12 (2012).
29. O'Hara, T., Virág, L., Varró, A. & Rudy, Y. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. *PLoS Comput. Biol.* **7**, e1002061 (2011).
30. Farrell, S. R., Ross, J. L. & Howlett, S. E. Sex differences in mechanisms of cardiac excitation-contraction coupling in rat ventricular myocytes. *Am. J. Physiol. Circ. Physiol.* **299**, H36–H45 (2010).
31. Parks, R. J. & Howlett, S. E. Sex differences in mechanisms of cardiac excitation–contraction coupling. *Pflügers Arch. - Eur. J. Physiol.* **465**, 747–763 (2013).
32. Parks, R. J., Ray, G., Bienvenu, L. A., Rose, R. A. & Howlett, S. E. Sex differences in SR Ca2+ release in murine ventricular myocytes are regulated by the cAMP/PKA pathway. *J. Mol. Cell. Cardiol.* **75**, 162–173 (2014).
33. Papp, R. *et al.* Genomic upregulation of cardiac Cav1.2α and NCX1 by estrogen in women. *Biol. Sex Differ.* **8**, 5–8 (2017).
34. Woosley, R., Heise, C., Gallo, T., Woosley, R. & Romero, K. QTDrug List. *AZCERT* Available at: https://crediblemeds.org/new-drug-list/.
35. Rush, S. & Larsen, H. A Practical Algorithm for Solving Dynamic Membrane Equations. *IEEE Trans. Biomed. Eng.* **BME-25**, 389–392 (1978).
36. Van Der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22 (2011).
37. McKinney, W. Data Structures for Statistical Computing in Python. in *Proceedings of the 9th Python in Science Conference* (eds. van der Walt, S. &

Millman, J.) 56–61 (2010). doi:10.25080/Majora-92bf1922-00a

38. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
39. Bergstra, J., Yamins, D. & Cox, D. D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28* I–115–I–123 (JMLR.org, 2013).
40. Verkerk, A. O. *et al.* Gender Disparities in Cardiac Cellular Electrophysiology and Arrhythmia Susceptibility in Human Failing Ventricular Myocytes. *Int. Heart J.* **46**, 1105–1118 (2005).
41. Fischer, T. H. *et al.* Sex-dependent alterations of Ca 2+ cycling in human cardiac hypertrophy and heart failure. *Europace* **18**, 1440–1448 (2016).
42. Furukawa, T. & Kurokawa, J. Regulation of cardiac ion channels via non-genomic action of sex steroid hormones: Implication for the gender difference in cardiac arrhythmias. *Pharmacol. Ther.* **115**, 106–115 (2007).
43. Kramer, J. *et al.* MICE Models: Superior to the HERG Model in Predicting Torsade de Pointes. *Sci. Rep.* **3**, 2100 (2013).
44. 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).
45. Yang, P.-C. *et al.* A computational model predicts adjunctive pharmacotherapy for cardiac safety via selective inhibition of the late cardiac Na current. *J. Mol. Cell. Cardiol.* **99**, 151–161 (2016).
46. Yang, P.-C. *et al.* A Computational Pipeline to Predict Cardiotoxicity. *Circ. Res.* **126**, 947–964 (2020).
47. Hondeghem, L. M., Carlsson, L. & Duker, G. Instability and Triangulation of the Action Potential Predict Serious Proarrhythmia, but Action Potential Duration Prolongation Is Antiarrhythmic. *Circulation* **103**, 2004–2013 (2001).
48. Parikh, J., Gurev, V. & Rice, J. J. Novel Two-Step Classifier for Torsades de Pointes Risk Stratification from Direct Features. *Front. Pharmacol.* **8**, 1–18 (2017).
49. Yang, P.-C. *et al.* A multiscale computational modelling approach predicts mechanisms of female sex risk in the setting of arousal-induced arrhythmias. *J. Physiol.* **595**, 4695–4723 (2017).
50. Gaborit, N. *et al.* Gender-related differences in ion-channel and transporter subunit expression in non-diseased human hearts. *J. Mol. Cell. Cardiol.* **49**, 639–646 (2010).
51. Rautaharju, P. M. *et al.* Sex differences in the evolution of the electrocardiographic QT interval with age. *Can. J. Cardiol.* **8**, 690–5 (1992).
52. Sims, C. *et al.* Sex, Age, and Regional Differences in L-Type Calcium Current Are Important Determinants of Arrhythmia Phenotype in Rabbit Hearts With Drug-Induced Long QT Type 2. *Circ. Res.* **102**, 86–100 (2008).
53. Tamari, I., Rabinowitz, B. & Neufeld, H. N. Torsade de pointes due to prenylamine controlled by lignocaine. *Eur. Heart J.* **3**, 389–92 (1982).
54. Champeroux, P. *et al.* Prediction of the risk of Torsade de Pointes using the model of isolated canine Purkinje fibres. *Br. J. Pharmacol.* **144**, 376–385 (2005).
55. Haverkamp, W. *et al.* Torsade de pointes induced by ajmaline. *Z. Kardiol.* **90**, 586–590 (2001).
56. Vieweg, W. V. R. *et al.* Proarrhythmic Risk with Antipsychotic and Antidepressant Drugs. *Drugs Aging* **26**, 997–1012 (2009).
57. Vieweg, W. V. R. *et al.* Risperidone, QTc interval prolongation, and torsade de pointes: A systematic review of case reports. *Psychopharmacology (Berl).* **228**,

515–524 (2013).
58. Harvey, P. A. & Leinwand, L. A. Oestrogen enhances cardiotoxicity induced by Sunitinib by regulation of drug transport and metabolism. *Cardiovasc. Res.* **107**, 66–77 (2015).
59. Wenzel-Seifert, K., Wittmann, M. & Haen, E. QTc Prolongation by Psychotropic Drugs and the Risk of Torsade de Pointes. *Dtsch. Aerzteblatt Online* **108**, 687–693 (2011).
60. Hasnain, M. *et al.* Quetiapine, QTc interval prolongation, and torsade de pointes : a review of case reports. *Ther. Adv. Psychopharmacol.* **4**, 130–138 (2014).
61. Darpo, B. *et al.* Are women more susceptible than men to drug-induced QT prolongation? Concentration-QT c modelling in a phase 1 study with oral rac-sotalol. *Br. J. Clin. Pharmacol.* **77**, 522–531 (2014).
62. Pokorney, S. D. *et al.* Dofetilide dose reductions and discontinuations in women compared with men. *Hear. Rhythm* **15**, 478–484 (2018).
63. Thompson, K. A., Murray, J. J., Blair, I. A., Woosley, R. L. & Roden, D. M. Plasma concentrations of quinidine, its major metabolites, and dihydroquinidine in patients with torsades de pointes. *Clin. Pharmacol. Ther.* **43**, 636–642 (1988).
64. Smetana, P., Batchvarov, V. N., Hnatkova, K., Camm, A. J. & Malik, M. Sex differences in repolarization homogeneity and its circadian pattern. *Am. J. Physiol. Circ. Physiol.* **282**, H1889–H1897 (2002).
65. Sredniawa, B., Musialik-Lydka, A., Jarski, P., Kalarus, Z. & Polonski, L. Circadian and sex-dependent QT dynamics. *Pacing Clin. Electrophysiol.* **28 Suppl 1**, S211-6 (2005).
66. Watanabe, J. *et al.* Increased Risk of Antipsychotic-Related QT Prolongation During Nighttime. *J. Clin. Psychopharmacol.* **32**, 18–22 (2012).
67. Täubel, J., Ferber, G., Fernandes, S. & Camm, A. J. Diurnal Profile of the QTc Interval Following Moxifloxacin Administration. *J. Clin. Pharmacol.* **59**, 35–44 (2019).
68. Lazzerini, P. E. *et al.* Proton Pump Inhibitors and Serum Magnesium Levels in Patients With Torsades de Pointes. *Front. Pharmacol.* **9**, 1–10 (2018).
69. Lv, C. *et al.* The Clinical Pharmacokinetics and Pharmacodynamics of Warfarin When Combined with Compound Danshen: A Case Study for Combined Treatment of Coronary Heart Diseases with Atrial Fibrillation. *Front. Pharmacol.* **8**, 1–10 (2017).
70. Huo, J., Wei, F., Cai, C., Lyn-Cook, B. & Pang, L. Sex-Related Differences in Drug-Induced QT Prolongation and Torsades de Pointes: A New Model System with Human iPSC-CMs. *Toxicol. Sci.* **167**, 360–374 (2018).

Tables

Table 1

Drug list with TdP risk category (TdP^+ in red, TdP^- in green, intermediate risk in yellow), IC_{50}s and EFTPC. All concentrations are expressed in nM. A Hill coefficient (nH) of 1 was used unless a different nH is indicated in parenthesis.

Drug	I_{Kr}	I_{NaL}	I_{CaL}	I_{Na}	I_{to}	I_{K1}	I_{Ks}	EFTPC
Amiodarone 1	860		1900	15900				0.8
Amiodarone 2	30		270	4800				0.5
Bepridil 1	160		1000	2300				35
Bepridil 2	33		211	3700				33
Bepridil CiPA	50 (0.9)	1813.9 (1.4)	2808.1 (0.6)	2929.3 (1.2)	8594 (3.5)	0 (0)	28628.3 (0.7)	33
Cilostazol	13800		91200	93700				128
Disopyramide	14400		1036700	168400				742
Dofetilide 1	30		26700	162100				2
Dofetilide 2	5		60000	300000				2
Dofetilide CiPA	4.9 (0.9)	753160.4 (0.3)	260.3 (1.2)	380.5 (0.9)	18.8 (0.8)	394.3 (0.8)	0 (0)	2
Donepezil	700		34300	38500				3
Flecainide	1500		27100	6200				753
Halofantrine	380		1900	331200				172
Haloperidol 1	40		1300	4300				4
Haloperidol 2	27		1700	7000				3.6
Ibutilide	18		62500	42500				140
Methadone	3500		37400	31800				507
Moxifloxacin	86200		173000	1112000				10960
Procainamide	272400		389500	746600				54180
Quinidine 1	720		6400	14600				3237
Quinidine 2	300		15600	16600				924
Quinidine CiPA	992 (0.8)	9417 (1.3)	51592.3 (0.6)	12329 (1.5)	3487.4 (1.3)	39589919 (0.4)	4898.9 (1.4)	3237
Sotalol	111400		193300	7013900				14690
Sotalol CiPA	110600 (0.8)	0 (0)	7061527 (0.9)	1140000000 (0.5)	43143455 (0.7)	3050260 (1.2)	4221856 (1.2)	14690
Sparfloxacin	22100		88800	2555000				1766
Terodiline	650		4800	7400				145
Thioridazine 1	500		3500	1400				980
Thioridazine 2	33		1300	1830				208
Ajmaline	1040		71000	8200				1500
Ceftriaxone	445700		153800	555900				23170
Cibenzoline	22600		30000	7800				976
Diazepam	53200		30500	306400				29
Diltiazem 1	13200		760	22400				122
Diltiazem 2	17300		450	9000				122

Diltiazem CiPA	13150 (0.9)	21868.5 (0.7)	112.1 (0.7)	110859 (0.7)	2820000000 (0.2)	0 (0)	0 (0)	122
Duloxetine	3800		2800	5100				16
Lamivudine	2054000		54200	1571400				19540
Linezolid	1147200		105400	2644500				59110
Loratadine	6100		11400	28900				0.4
Mexiletine	50000		100000	43000				4129
Mexiletine CiPA	28880 (0.9)	8956.8 (1.4)	38243.6	0 (0)	0 (0)	0 (0)	0 (0)	4129
Mibefradil1	1700		510	5600				12
Mibefradil2	1800		156	980				12
Mitoxantrone	539400		22500	93500				225
Nifedipine1	44000		12	88500				8
Nifedipine2	275000		60	37000				7.7
Nitrendipine1	24600		25	21600				3
Nitrendipine2	10000		0.35	36000				3.02
Pentobarbital	1433900		299000	2686000				5171
Phenytoin1	147000		21900	72400				4360
Phenytoin2	100000		103000	49000				4500
Prenylamine	65		1240	2520				17
Propranolol	2828		18000	2100				26
Ribavirin	967000		622500	2997500				27880
Sitagliptin	174700		147100	1220800				442
Telbivudine	422700		713900	1095200				19720
Verapamil1	250		200	32500				88
Verapamil2	143		100	41500				81
Verapamil_CiPA	288	7028	201.8 (1.1)	0 (0)	13429.2 (0.8)	349000000 (0.3)	0 (0)	81
Amitriptyline	3280		11600	20000				41
Astemizole	4		1100	3000				0.3
Chlorpromazine1	1500		3400	3000				38
Chlorpromazine_CiPA	929.2 (0.8)	4559.6 (0.9)	8191.9 (0.8)	4535.6 (2)	17616711 (0.4)	9269.9 (0.7)	0 (0)	38
Cisapride1	20		11800	337000				3
Cisapride_CiPA	10.1 (0.7)	0 (0)	9258076 (0.4)	0 (0)	219112.4 (0.2)	29498 (0.5)	81192862 (0.3)	2.6
Clozapine	2300		3600	15100				71
Dasatinib	24500		81100	76300				41
Desipramine	1390		1709	1520				108
Diphenhydramine	5200		228000	41000				34
Droperidol	60		7600	22700				16
Fluvoxamine	3100		4900	39400				377
Imipramine	3400		8300	3600				106
Metronidazole	1340200		177900	2073200				187000
Nilotinib	1000		17500	13300				172
Ondansetron_CiPA	1320 (0.9)	19180.8	22551.4 (0.8)	57666.4	1023378	0 (0)	569807 (0.7)	139
Paliperidone	780		193900	109000				69

Paroxetine	1900		3900	9800				14
Pimozide1	40		240	1100				0.5
Pimozide2	20		162	54				1
Piperacillin	3405100		1226000	2433800				1378000
Propafenone	440		1800	1190				241
Quetiapine	5800		10400	16900				33
Raltegravir	782800		246700	824200				7000
Ranolazine_CiPA	8270 (0.9)	7884.5 (0.9)	0 (0)	68774 (1.4)	0 (0)	0 (0)	36155020 (0.5)	1948.2
Risperidone1	260		34200	43400				2
Risperidone2	150		73000	102000				1.81
Saquinavir	16900		1900	12100				130
Sertindole1	33		6300	6900				2
Sertindole2	14		8900	2300				1.59
Solifenacin	280		4300	1500				3
Sunitinib	1200		33400	16500				13
Terfenadine1	50		930	2000				9
Terfenadine2	8.9		375	971				9
Terfenadine_CiPA	23 (0.6)	20056 (0.6)	700.4	4803.2	239960.8 (0.3)	0 (0)	399754 (0.5)	4
Voriconazole	490900		414200	1550500				7563

Table 2

List of biomarkers extracted from simulations and relative descriptions.

Biomarker	Description
APD_{90}	*Action potential duration at 90% repolarization*
APD_{75}	*Action potential duration at 75% repolarization*
APD_{50}	*Action potential duration at 50% repolarization*
APD_{30}	*Action potential duration at 30% repolarization*
V_{max}	*Peak voltage*
V_{min}	*Diastolic voltage*
$dVdt_{max}$	*Maximal upstroke velocity*
Plateau potential	*Average voltage between 10 and 50 ms after action potential initiation*
I_{Na} max	*Peak of the sodium current*
CaT max	*Peak concentration of the calcium transient*
CaT min	*Diastolic intracellular calcium concentration*
CaD_{80}	*Calcium transient duration at 80% return to baseline*
CaT tau	*Rate constant of decay of the calcium transient*
Ca_i integral	*Integral of calcium transient*
Na_i integral	*Integral of sodium transient ???*
I_{Na} integral	*Integral of fast component of the sodium current*
I_{NaL} integral	*Integral of late component of the sodium current*
I_{to} integral	*Integral of transient outward potassium current*
I_{CaL} integral	*Integral of L-type calcium current*
I_{Kr} integral	*Integral of rapid delayed rectifier potassium current*
I_{Ks} integral	*Integral of slow delayed rectifier potassium current*
I_{K1} integral	*Integral of inward rectifier potassium current*
I_{NaK} integral	*Integral of sodium/potassium-ATPase current*
I_{pCa} integral	*Integral of calcium pump current*
I_{NaCa} integral	*Integral of sodium/calcium exchanger current*
AP amplitude	*Action potential amplitude*
AP triangulation	*Difference between APD_{80} and APD_{30}*

Table 3

Summary of prediction performances of the different combination of TdP classifiers and dataset discussed.

	F1 score	MCC	AUC	FN	FP	Misclassified drugs
M on M	0.9148	0.8336	0.9447	4	1	Amiodarone 1, Amiodarone 2, Cilostazol, Donepezil, Prenylamine
F on F	0.9492	0.8987	0.9424	1	2	Ajmaline, Prenylamine, Procainamide
M on F	0.7693	0.5945	0.9159	12	1	Amiodarone 1, Amiodarone 2, Bepridil 2, Bepridil CiPA, Cilostazol, Donepezil, Flecainide, Halofantrine, Prenylamine, Procainamide, Quinidine 1, Sotalol, Thioridazine 1
M re-trained on F	0.8980	0.7972	0.9217	4	2	Ajmaline, Amiodarone 1, Amiodarone 2, Cilostazol, Donepezil, Prenylamine
F on EF	0.9322	0.8665	0.9516	1	3	Ajmaline, Prenylamine, Procainamide, Telbivudine
F on LF	0.9492	0.8987	0.9516	1	2	Ajmaline, Prenylamine, Procainamide
F on LU	0.9492	0.8987	0.9505	1	2	Ajmaline, Prenylamine, Procainamide