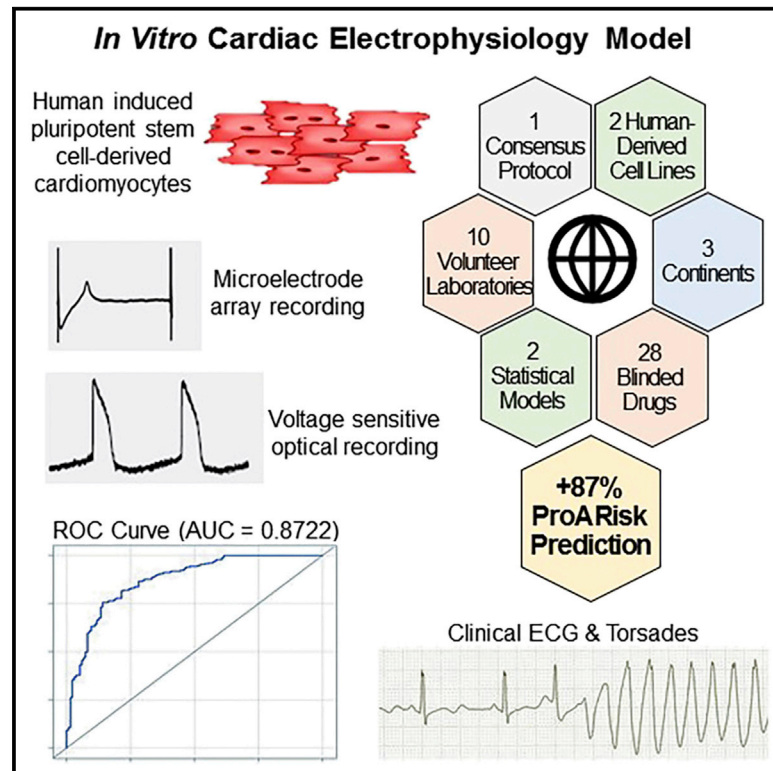


Cell Reports

International Multisite Study of Human-Induced Pluripotent Stem Cell-Derived Cardiomyocytes for Drug Proarrhythmic Potential Assessment

Graphical Abstract



Authors

Ksenia Blinova, Qianyu Dang, Daniel Millard, ..., Norman Stockbridge, David G. Strauss, Gary Gintant

Correspondence

ksenia.blinova@fda.hhs.gov (K.B.), gary.gintant@abbvie.com (G.G.)

In Brief

Blinova et al. tested human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) for improving torsades de pointes arrhythmia risk prediction of drugs in the Comprehensive *In Vitro* Proarrhythmia Assay (CiPA) initiative. This validation study confirms their utility based on electrophysiologic responses to 28 blinded drugs, with minimal influence from cell lines, test sites, and electrophysiological platforms.

Highlights

- 28 drugs classified by torsadesproarrhythmic risk studied in blinded experiments
- Two commercial human cardiomyocyte lines used with 5 devices across 10 sites
- Repolarization effects evaluated using microelectrode array and voltage-sensing dyes
- Statistical model predicted proarrhythmic risk from electrophysiologic responses



International Multisite Study of Human-Induced Pluripotent Stem Cell-Derived Cardiomyocytes for Drug Proarrhythmic Potential Assessment

Ksenia Blinova,^{1,24,*} Qianyu Dang,² Daniel Millard,³ Godfrey Smith,^{4,5} Jennifer Pierson,⁶ Liang Guo,⁷ Mathew Brock,⁸ Hua Rong Lu,⁹ Udo Kraushaar,¹⁰ Haoyu Zeng,¹¹ Hong Shi,¹² Xiaoyu Zhang,¹³ Kohei Sawada,^{14,15} Tomoharu Osada,¹⁶ Yasunari Kanda,¹⁷ Yuko Sekino,^{15,17} Li Pang,¹⁸ Tromondae K. Feaster,¹⁹ Ralf Kettenhofen,²⁰ Norman Stockbridge,²¹ David G. Strauss,²² and Gary Gintant^{23,*}

¹Division of Biomedical Physics, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, MD 20993, USA

²Office of Biostatistics, Office of Translational Science, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20993, USA

³Axion BioSystems, Atlanta, GA 30309, USA

⁴University of Glasgow, Glasgow G12 8QQ, Scotland, UK

⁵Clyde Biosciences, Newhouse ML1 5UH, Scotland, UK

⁶Health and Environmental Sciences Institute, Washington, DC 20005, USA

⁷Investigative Toxicology, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Frederick, MD 21702, USA

⁸Genentech, San Francisco, CA 94080, USA

⁹Discovery Sciences, R&D, Janssen Pharmaceutical (JNJ), Beerse, Belgium

¹⁰NMI Natural and Medical Sciences Institute at the University of Tübingen, Reutlingen, Germany

¹¹Merck, Safety & Exploratory Pharmacology Department, West Point, PA 19486, USA

¹²Bristol-Myers Squibb, New York, NY 10154, USA

¹³ACEA Biosciences, San Diego, CA 92121, USA

¹⁴Eisai, Tsukuba, Ibaraki 300-2635, Japan

¹⁵The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

¹⁶LSI Medience, Chiyoda-ku, Tokyo 101-8517, Japan

¹⁷Division of Pharmacology, National Institute of Health Sciences, Kawasaki, Kanagawa 210-9501, Japan

¹⁸Division of Biochemical Toxicology, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079, USA

¹⁹Cellular Dynamics International—A FUJIFILM Company, Madison, WI 53711, USA

²⁰Ncardia, Cologne 50829, Germany

²¹Division of Cardiovascular and Renal Products, Office of Drug Evaluation I, Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20993, USA

²²Division of Applied and Regulatory Science, Office of Clinical Pharmacology, Office of Translational Science, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20993, USA

²³AbbVie, North Chicago, IL 60064-6118, USA

²⁴Lead Contact

*Correspondence: ksenia.blinova@fda.hhs.gov (K.B.), gary.gintant@abbvie.com (G.G.)

<https://doi.org/10.1016/j.celrep.2018.08.079>

SUMMARY

To assess the utility of human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) as an *in vitro* proarrhythmia model, we evaluated the concentration dependence and sources of variability of electrophysiologic responses to 28 drugs linked to low, intermediate, and high torsades de pointes (TdP) risk categories using two commercial cell lines and standardized protocols in a blinded multisite study using multielectrode array or voltage-sensing optical approaches. Logistical and ordinal linear regression models were constructed using drug responses as predictors and TdP risk categories as outcomes. Three of seven predictors (drug-induced arrhythmia-like events and prolongation of repo-

larization at either maximum tested or maximal clinical exposures) categorized drugs with reasonable accuracy (area under the curve values of receiver operator curves ~ 0.8). hiPSC-CM line, test site, and platform had minimal influence on drug categorization. These results demonstrate the utility of hiPSC-CMs to detect drug-induced proarrhythmic effects as part of the evolving Comprehensive *In Vitro* Proarrhythmia Assay paradigm.

INTRODUCTION

Fourteen drugs have been removed from the market worldwide as a result of their potential to induce a rare but potentially fatal ventricular arrhythmia, torsades de pointes (TdP) (Stockbridge et al., 2013). The International Council on Harmonisation (ICH)



adopted two guidelines on the assessment of drug-induced TdP (ICH S7B and ICH E14) that outline the assessment of the potential of new pharmaceuticals to delay ventricular repolarization in *in vitro* assays, including testing for their ability to block the human ether-a-go-go-related (hERG) potassium channel, and *in vivo*, to prolong the QT interval on the electrocardiogram. Adoption of these guidelines has been effective in preventing new drugs with unrecognized TdP risk from reaching the market; however, the current regulatory approach lacks specificity, because multiple drugs block hERG or prolong the QT interval but have a low risk of TdP. It is possible that overemphasis on hERG block and QT prolongation in proarrhythmic potential assessment has prevented some useful and safe drugs from reaching the market. The Comprehensive *In Vitro* Proarrhythmia Assay (CiPA) initiative represents a new paradigm to improve the specificity of proarrhythmic risk assessment (Fermini et al., 2016; Sager et al., 2014). The non-clinical aspects of CiPA rely on a mechanistic assessment of drug effects on cellular electrophysiology (EP) using (1) *in silico* reconstruction of human ventricular electrical activity based on drug effects on multiple human ionic currents, each expressed in heterologous expression systems, and (2) assessment of drug effects in human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) to detect any missed or unanticipated EP effects (Food and Drug Administration, 2017).

The use of hiPSC-CMs for cardiac safety evaluation of the new drug candidates continues to increase, as evidenced by numerous recent publications. Many of these studies demonstrate the ability of hiPSC-CMs as model systems to detect EP effects of drugs, including delayed or altered repolarization (Blinova et al., 2017; Clements and Thomas, 2014; Yamamoto et al., 2016). While encouraging, such studies typically use small test sets; different cellular preparations, protocols, and experimental endpoints; inconsistent criteria to interpret results; and different gold standards related to either delayed repolarization or proarrhythmic risk. Such differences hinder cross-site comparisons of data and recognition of sources of experimental variability. A significant step forward was made recently (Ando et al., 2017; Yamamoto et al., 2016), in which a large set of drugs was evaluated at multiple sites following a standardized experimental protocol; however, that study was limited to the evaluation of a single cell line and one EP platform used across sites with no statistical modeling of results. Comprehensive evaluations using multiple sites, interrogation techniques, and cell sources are necessary because all models have limitations that may appear under different circumstances. Despite possessing nearly identical underlying early after depolarization (EAD) properties as traditionally accepted models (e.g., mature canine ventricular cardiomyocytes [Ma et al., 2011]), hiPSC-CMs are often described as having fetal or neonatal ion channel and ionic current stoichiometries (Jonsson et al., 2012; Sala et al., 2017) that may interfere with the accurate prediction of proarrhythmic risk.

To characterize the potential utility of hiPSC-CMs within the CiPA paradigm, the present study was conducted to characterize, in blinded fashion, the EP effects of 28 drugs with known clinical TdP risk on hiPSC-CMs using 2 commercially available hiPSC-CM lines tested across 10 experimental sites and 5 EP

platforms. Specifically, this validation study focused on (1) characterization of site-to-site variability of the assessment of EP effects of the drugs using either microelectrode array (MEA) or voltage-sensing optical (VSO) techniques and standardized protocols to assess drug-induced altered repolarization, and (2) identification of important hiPSC-CM assay endpoints associated with high, intermediate, and low TdP risk using linear regression models. The present study builds upon on a previous smaller pilot study that evaluated the EP effects of 8 drugs using MEA approaches and 4 positive controls across a smaller number of sites (Millard et al., 2018). Overall, the conceptual advance of this work is not in the discovery or the development of a new iPSC-CMs-based assay, but rather in performing a first large-scale multisite study combining MEA and VSO techniques to evaluate the current state of iPSC-CM-based assays in the assessment of drug-induced TdP.

RESULTS

Electrophysiological Effects Induced in hiPSC-CMs by Drugs with Known Risk Levels for Clinical TdP Risk

Ten independent sites used a standardized protocol to evaluate the EP effects of 28 drugs with known levels of clinical risk categorized by expert consensus (Colatsky et al., 2016) into 3 clinical TdP risk groups (high, intermediate, and low or no risk) (Fermini et al., 2016) (see Table S1 for details on experimental sites, protocols, and platforms and Table S2 for drug categories and concentrations). Fifteen complete datasets, including data from 5 sites that studied drug effects on both iCell cardiomyocytes² and Cor.4U cardiomyocytes (referred to here as iCell² and Cor.4U cells), were analyzed. Drug-induced repolarization prolongation (baseline and vehicle-controlled, rate-corrected action potential duration at 90% repolarization, ddAPD90c, or field potential duration, ddFPDc, depending on the EP platform used; see Experimental Procedures) was measured along with drug-induced arrhythmias of different types (Figure 1) or drug-induced cessation of hiPSC-CMs' spontaneous beating (i.e., quiescence).

Low TdP Risk Category

The 9 drugs in the low TdP risk category were verapamil, diltiazem, loratadine, metoprolol, mexiletine, nifedipine, nitrendipine, ranolazine, and tamoxifen. hiPSC-CM responses to verapamil for all 10 sites is shown in Figure 2, and corresponding figures for the rest of the low-risk drugs are shown in Supplemental Data S1. No verapamil-induced repolarization prolongation or arrhythmias were observed at any concentrations studied.

Diltiazem, loratadine, nifedipine, nitrendipine, and tamoxifen did not induce any arrhythmias or statistically significant repolarization prolongation at concentrations up to 20- to 140-fold clinical C_{max}. The remaining 3 drugs (ranolazine, metoprolol, and mexiletine) induced repolarization prolongation and arrhythmias at ≥ 1 of the concentrations in several datasets.

Ranolazine is a known hERG blocker (Crumb et al., 2016) at clinical concentrations and produces QT prolongation. Consistent with this, 13 of 15 datasets show statistically significant ranolazine-induced repolarization prolongation at concentrations between 0.1- and 5.0-fold C_{max}, but no

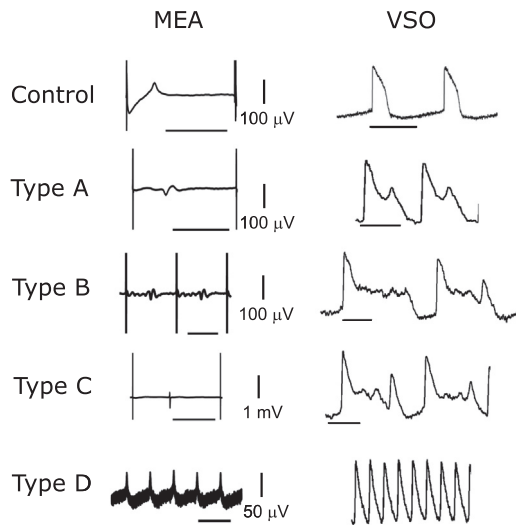


Figure 1. Representative Traces of Four Cellular Arrhythmia-Like Events Recorded in hiPSC-CMs

Recorded by (left) MEA and (right) VSO platforms. The horizontal scale bar equals 1 s. We refer to type A arrhythmia as a “mild” arrhythmia-like event in the text.

ranolazine-induced arrhythmias. At the highest studied concentration, 100 μM , or >50-fold C_{max} , 6 of 15 datasets show ranolazine-induced arrhythmia-like events or cessation of spontaneous beating (9 of 15 datasets) in at least 2 experimental wells.

Metoprolol is a beta-1 blocker that slows heart rate clinically and is not associated with TdP risk. In hiPSC-CMs, where beta-blockade does not occur due to the absence of sympathetic innervation, metoprolol-induced arrhythmias occurred at 100 μM (55-fold C_{max}) in 5 datasets and at 31.6 μM (~18-fold C_{max}) in 1 dataset, which is consistent with metoprolol-induced hERG block at higher concentrations (drug concentration for 50% block [IC_{50}] = 145 μM) (Kawakami et al., 2006). Finally, 10 μM mexiletine (~4-fold C_{max}) induced arrhythmias in 3 of 15 datasets and mexiletine-induced cessation of spontaneous beating at the highest concentration (100 μM , ~40-fold C_{max}) in 12 of 15 datasets.

Intermediate TdP Risk Category

The 11 drugs in the intermediate TdP risk category were terfenadine, astemizole, chlorpromazine, cisapride, clarithromycin, clozapine, domperidone, droperidol, ondansetron, pimozone, and risperidone. hiPSC-CM response to terfenadine for all 10 sites is shown in Figure 3, and corresponding figures for the other intermediate risk drugs are shown in Supplemental Data S2. None of the sites observed terfenadine-induced arrhythmia-like events in hiPSC-CMs, even at concentrations as high as 350-fold C_{max} , but terfenadine-induced repolarization prolongation occurred in 11 of 15 datasets.

Statistically significant repolarization prolongation at ≥ 1 studied concentrations was observed in a minimum of 10 of 15 datasets for all of the drugs in the intermediate-risk category but clozapine and chlorpromazine. Clozapine- and chlorpromazine-

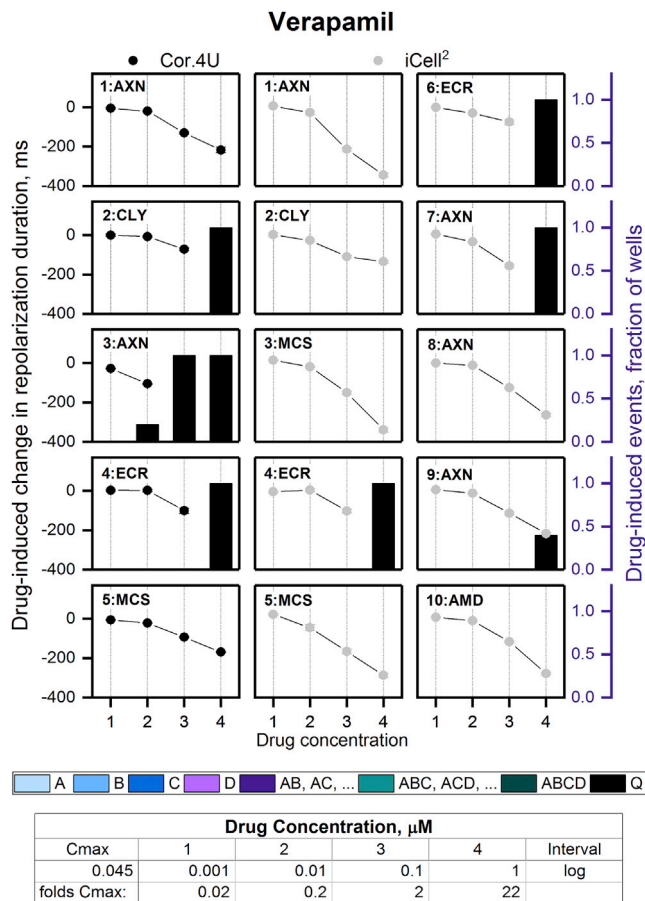


Figure 2. EP Effects of Verapamil (Low TdP Risk) across 10 Sites (15 Site/Cell Combinations)

Panel titles represent site number followed by a three-letter code of EP platform used (AXN, Maestro [Axion BioSystems]; CLY, CellOPTIQ [Clyde Biosciences]; ECR, CardioECR [ACEA Biosciences]; AMD, AlphaMED64 [Alpha MED Scientific]; and MCS, MEA2100 [Multichannel Systems]). Drug-induced repolarization prolongation (black and gray circles for Cor.4U and iCell², correspondingly, left y axis) are shown as averaged baseline- and vehicle-controlled, Fridericia rate-corrected $\text{ddFPDc}/\text{ddAPD90c}$. Error bars represent SEs. The bars represent the percentage of wells in which a particular arrhythmic or quiescent event was observed (see color legend). $\text{ddFPDc}/\text{APD90c}$ was not calculated for the drug concentrations in which $\geq 50\%$ of the wells included in the analysis were arrhythmic after drug addition. Drug concentrations (in μM and x-fold above free [unbound] clinical C_{max} values) are shown in the table on the bottom of the figure, along with the concentration intervals.

See also Data S1.

zine-induced prolongation was reported in only 1 and 3 of 15 of the datasets, respectively. Drug-induced arrhythmia-like events at any concentration were observed in at least 10 of 15 datasets for all of the intermediate-risk drugs, except for chlorpromazine, clozapine, terfenadine, and risperidone (0–2 datasets of 15 contained arrhythmia events for these 4 drugs).

High TdP Risk Category

The 8 drugs in the high TdP risk category were dofetilide, azimilide, bepridil, D_{1L} -sotalol, disopyramide, ibutilide, quinidine, and

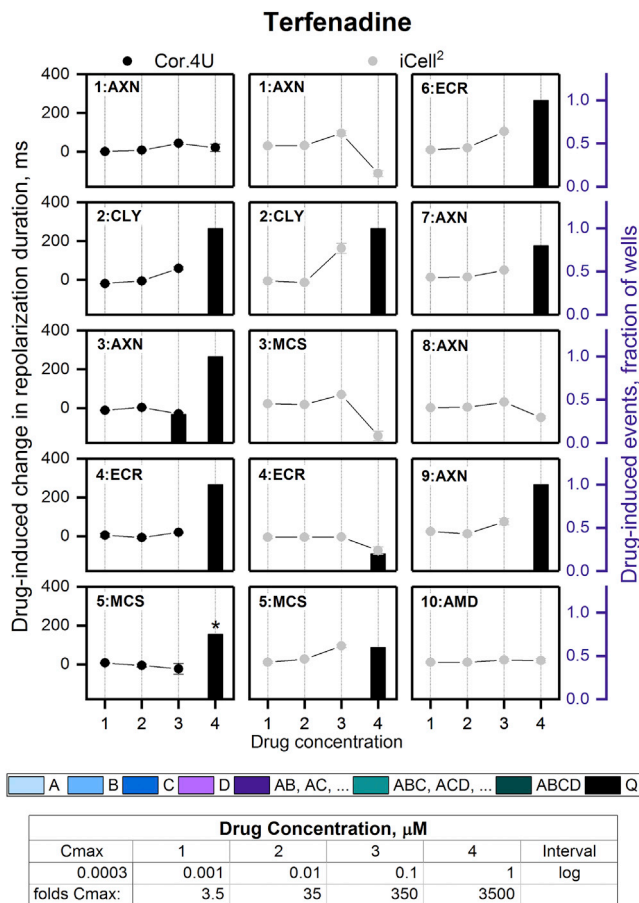


Figure 3. EP Effects of Terfenadine (Intermediate TdP Risk) across 10 Sites (15 Site/Cell Combinations)
See Figure 2 legend. A star represents data with number of replicate wells $N < 5$. See also Data S2.

vandetanib. The hiPSC-CM response to dofetilide for all 10 sites is shown in Figure 4, and corresponding figures for the other high-risk drugs are shown in Supplemental Data S3. Statistically significant dofetilide-induced repolarization prolongation or dofetilide-induced arrhythmia-like events were consistently (14 of 15 datasets) observed in the studied drug concentration range (0.16- to 5-fold Cmax).

All of the drugs in this category except bepridil induced statistically significant repolarization prolongation and/or arrhythmia-like events in both hiPSC-CM lines in at least 10 of 15 datasets. While bepridil-induced statistically significant repolarization prolongation was reported in 8 of 15 datasets, only 2 datasets contained bepridil-induced arrhythmia-like events. Drug-induced arrhythmia-like events were consistently observed at concentrations close to clinical Cmax for dofetilide, quinidine, and D,L-sotalol and at concentrations well below Cmax for ibutilide. Some of the drugs in this category were so potent and the chosen concentration escalation rate was so steep (i.e., logarithmic increase) that there were no detectable drug effects at one of the studied concentrations, and then at the next concentration, all of the hiPSC-CMs demonstrated arrhythmia-like

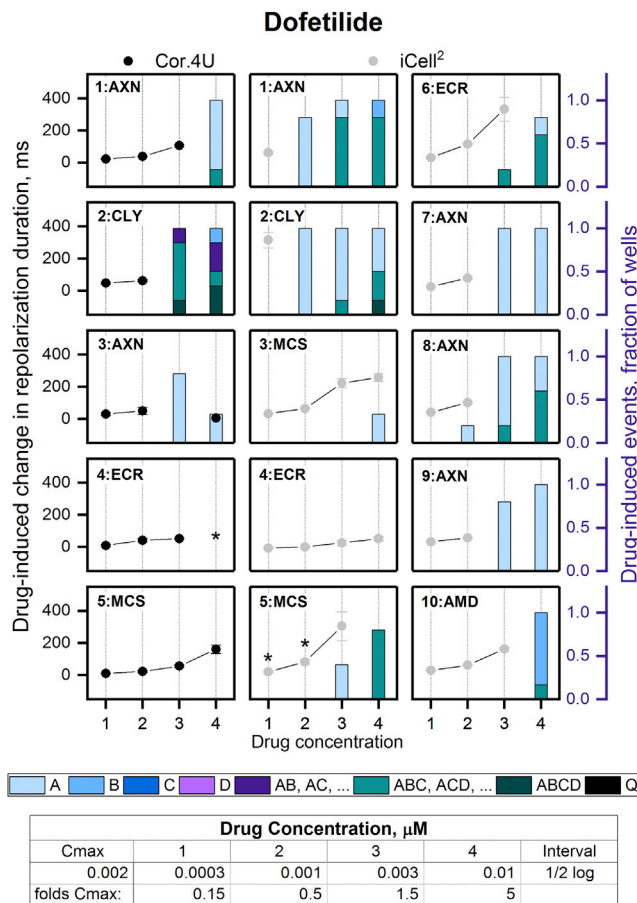


Figure 4. Effects of Dofetilide (High TdP Risk) across 10 Sites (15 Site/Cell Combinations)
See Figure 2 legend. Stars represent missing data or data with number of replicate wells $N \leq 5$. See also Data S3.

events, preventing reliable measurement of repolarization duration.

Minimal Effect of Site-to-Site Variability on Drug-Induced ddFPDc/APD90c

Despite significant efforts to apply consistent experimental protocols across sites, minor deviations were noted (see Table S1 for the experimental protocol deviations for each site). Site-to-site variability in drug-induced ddFPDc/APD90c averaged across all 28 drugs was compared to other sources of variability by treating site effects as either fixed or random effects (Table S3) and using the square root of the mean squared error (SR MSE) for each contribution. When site effect was treated as a fixed effect, SR MSE introduced by site (170 ms) was lower than variability induced by the hiPSC-CM line (245 ms). As expected, both values were lower than the contribution provided by drug concentration (482 ms). Similarly, if site effects were treated as random effects, the variability in drug-induced ddFPDc or ddAPD90c averaged over 28 drugs introduced by the site was lower than the total

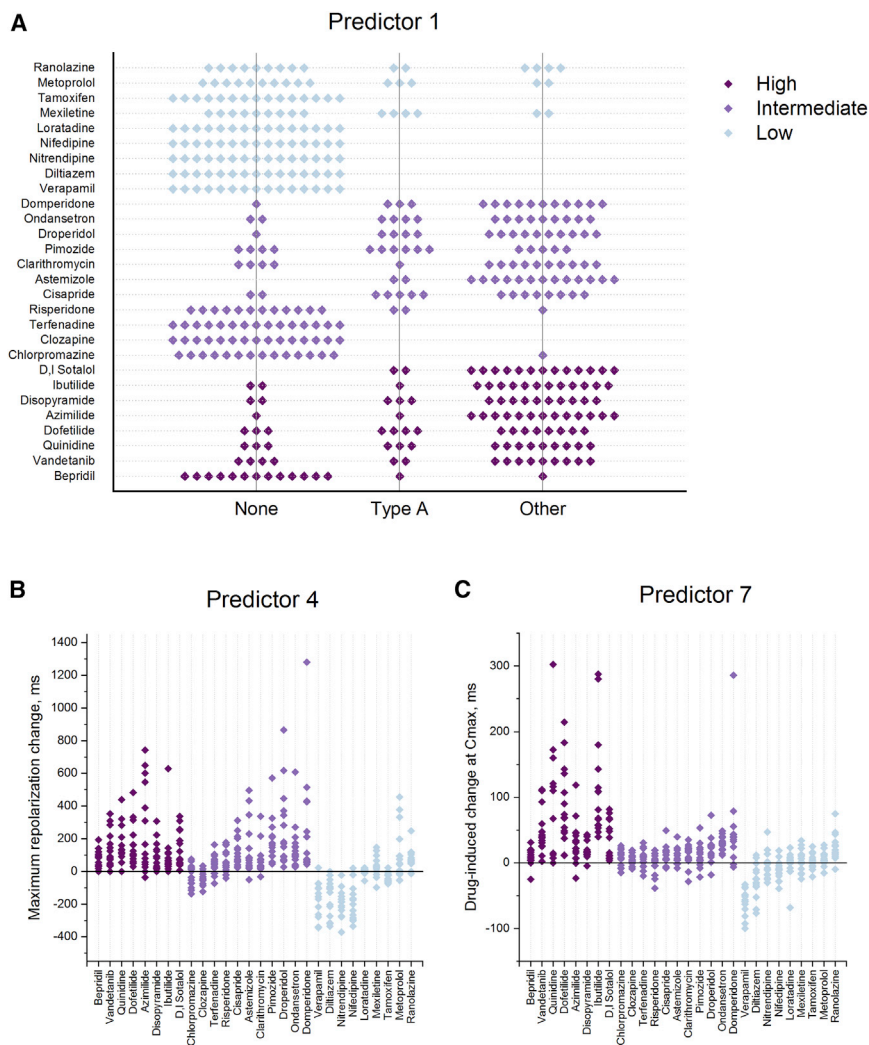


Figure 5. Three Significant Model Predictors for Model 1 Shown for All 28 Drugs

Each data point represents individual dataset (site/cell type combination, 15 datasets total). (A) Predictor 1, drug-induced arrhythmia-like event at any concentration (none, no arrhythmias; type A, only arrhythmia type A; other, any other arrhythmia type: B, C, D, or any combination of ≥ 2 arrhythmia types). (B) Predictor 4, maximum observed drug-induced repolarization prolongation or shortening (ddFPDc or ddAPD90c) at all studied drug concentrations. (C) Predictor 7, estimated drug-induced repolarization prolongation or shortening (ddFPDc or ddAPD90c) at clinical Cmax.

event (predictor 6) were first observed; predictor 7 is an estimated amount of prolongation that a drug would induce at the clinical Cmax (Experimental Procedures).

Logistic regression models were used in the regression of risk group (high or intermediate risk versus low risk [model 1], and high risk versus low risk and intermediate risk versus low risk [model 2]) on all 7 risk predictors. Cluster analysis showed that the pairs of predictors 3 and 4 (Pearson correlation = 0.52) and predictors 5 and 6 (Pearson correlation = 0.65) are highly correlated, so 1 of each pair may be redundant (data not shown). The final fitted models included 3 significant predictors: predictor 1, predictor 4, and predictor 7. Figure 5 shows significant model predictors for all of the sites for each drug. hiPSC-CM type (iCell² or Cor.4U)

random variability from all of the other sources of random variability (36 versus 67 ms), including well-to-well variability, plate-to-plate variability, human error, and other sources of variability.

Modeling of Drug Proarrhythmic Potential Based on Its hiPSC-CM Effects

Data on the EP effects of 28 drugs with a known clinical risk of TdP obtained from all of the experimental sites were used to construct a model that would predict TdP risk category of a drug based on its effects on hiPSC-CMs. Seven endpoints from hiPSC-CM experiments were used as potential model predictors (Table S4). Predictors 1 and 2 describe the ability of a drug to induce arrhythmia-like events in hiPSC-CMs; predictors 3 and 4 reflect the amount of drug-induced repolarization prolongation (ddFPDc or ddAPD90c) at the lowest concentration at which statistically significant change from the baseline (predictor 3) or maximum prolongation at any of the studied concentrations (predictor 4) was observed; predictors 5 and 6 account for concentrations of a drug relative to its clinical Cmax when prolongation of FP/AP duration (predictor 5) or arrhythmia-like

was not significant ($p = 0.089$) and did not improve overall fitting for model 1, but it showed a slight improvement in fitting for model 2 by decreasing the Akaike information criterion (AIC) value from 705.3 to 703.2, where AIC is an estimator or the relative quality of statistical models (for a given set of data, smaller value indicates better fit):

$$\text{Logit}(P1) = (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) \quad (1)$$

$$\text{Logit}(P2a) = (\text{Cell Type}) + (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) \quad (2)$$

$$\text{Logit}(P2b) = (\text{Cell Type}) + (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) \quad (3)$$

where $\text{Logit}(P) = \log(P/(1-P))$, P1 is a probability of a drug to be of high or intermediate TdP risk in model 1 and P2a and P2b are probabilities of a drug to be of high versus low or intermediate versus low TdP risk in model 2, respectively. Detailed model parameters are shown in Table S5. Averaged across sites, risk

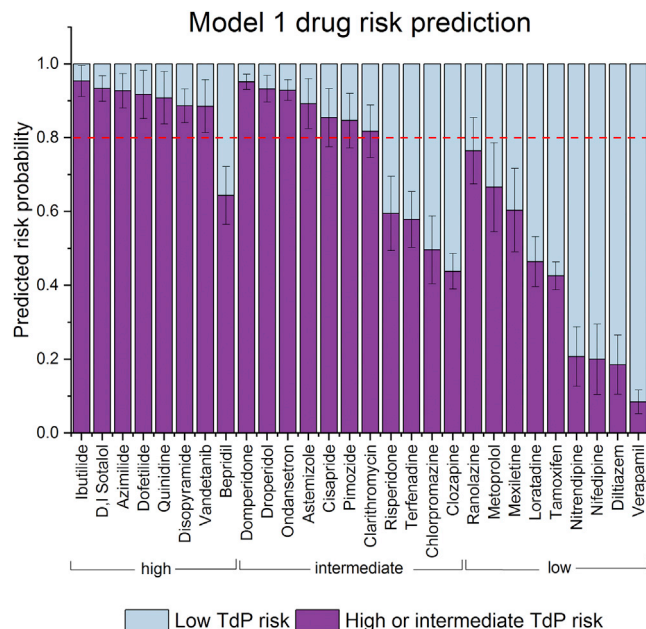


Figure 6. Model 1 (Dichotomous Model) Prediction of a Drug's TdP Risk Category to Be Either Low or Intermediate and High Combined Averaged from All 10 Sites (15 Cell Type/Platform Combinations)

Error bars represent 95% confidence intervals. Red dotted line represents the 0.8 threshold discussed in the text.

probabilities predicted by models 1 and 2 are shown in Figures 6 and 7, respectively. An example of model 1 and model 2 prediction is provided in Figure S1. Model 1 prediction fitted through the data of all of the sites had an area under the receiver operating characteristic (ROC) curve (AUC) value of 0.872 (Figure S2). As expected for a model with 3 outcomes, model 2 AUC had a lower value of 0.826 (Figure S2). Concordance indices (Somers' deltas [Somers' D], a measure of ordinal association between possibly dependent random variables, values from -1 to 1 , with higher values indicating better quality of model prediction) calculated for models 1 and 2 were 0.74 and 0.65, respectively, showing good discriminating utility for both models.

Figure 6 can be used to illustrate the potential role of hiPSC-CMs as a high-specificity preclinical assay under CiPA. By setting a threshold of low TdP risk versus high or intermediate TdP risk at 0.8 in model 1 (Figure 6), the predicted TdP risk of all of the drugs in the low-risk category fall below the threshold, providing a user with reasonable confidence that no unanticipated effects were missed for a drug. The one exception is ranolazine, for which the upper confidence interval (CI) of its estimated risk crosses the 0.8 threshold. As has been shown before (Blinova et al., 2017), the TdP risk of drugs that have significant late sodium current effects (e.g., ranolazine) may not be adequately modeled by existing hiPSC-CMs. Model 1 risk prediction fell below the 0.8 threshold for 1 drug from the high TdP risk category (bepiridil) and 4 drugs from the intermediate drug risk category (risperidone, terfenadine, chlorpromazine, and clozapine), highlighting the limitations of the current hiPSC-CMs assays that are not developed to be used as a stand-alone assay, but can be useful when

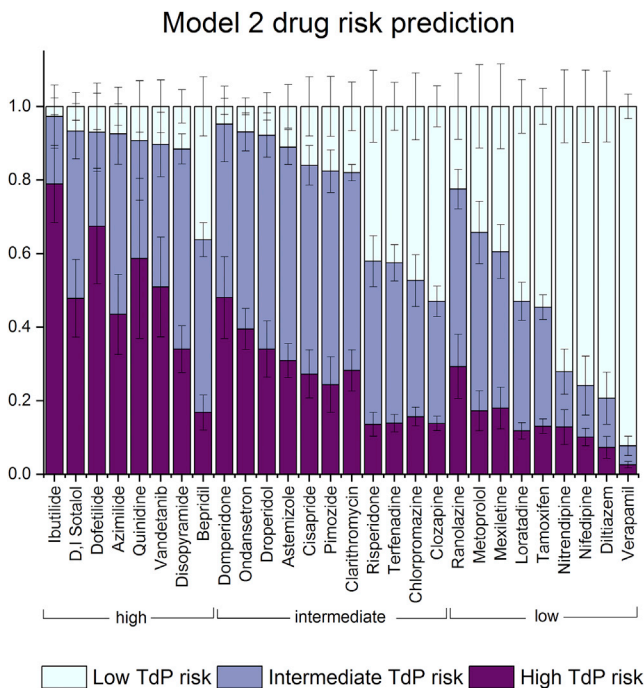


Figure 7. Model 2 Prediction of a Drug to Fall into Low, Intermediate, or High TdP Risk Category Averaged across 10 Sites (15 Cell Type/Platform Combinations)

Error bars represent 95% confidence intervals.

combined with other CiPA preclinical proarrhythmia assessment strategies.

Model Validation

The purpose of model validation is to estimate the performance of a model when applied for a new, independent dataset. One approach would be to split the data into training and validation sets, so one may use the training dataset to develop the model and then apply the model to the validation dataset to measure the performance. This approach usually requires a large sample size to avoid significant power loss for modeling. In the present study, due to the limited sample size, we have performed model validation and calibration using two alternative methods: cross-validation and bootstrapping. Both approaches allow for nearly unbiased estimates of future model performance, assuming that the present study sample represents a true random sampling of the population of interest. For cross-validation, the original data are randomly divided into k equally sized subsamples, then one subsample is used as the validation dataset, while the remaining subsamples are used as training data. The cross-validation process is then repeated k times for each subset. The k results then are averaged to produce a single estimation. Here, $k = 10$ was used for the cross-validation process. Similarly, bootstrapping uses re-sampling with the replacement from the original dataset, so theoretically, an infinite number of samples from one set of data can be generated. For both methods, if the analysis from re-sampling produces results that are consistent with the original analysis, then the model is

considered to be reliable and expected to perform for a new, independent dataset. For model 1, bootstrapping with 500 runs of re-sampling negligibly reduced the AUC, from 0.872 to 0.865. Similarly, cross-validation of model 1, omitting from the model 10% of all observations at a time, minimally reduced the AUC, from 0.872 to 0.862. Both results demonstrate the high reliability of model 1. For model 2, bootstrapping with 500 runs of re-sampling negligibly reduced the AUC value, from 0.819 to 0.808, demonstrating a robust model. Model validation results suggest that they would be practical, even when used by a single site applying one of the tested EP platform and cell type combinations. The experimental design used in the statistical models does not provide sufficient power to evaluate differences in the performance of the EP platforms or iPSC-CMs lines.

DISCUSSION

This study summarizes results of the first large multisite study assessing the potential of 2 commercially available hiPSC-CMs using *in vitro*-based MEA and VSO approaches to detect drug-induced repolarization abnormalities and predict the proarrhythmic potential of 28 drugs characterized for TdP risk under the CiPA initiative. Concentration-dependent effects from 7 EP responses were used to build 2 regression models that predict low, intermediate, or high clinical TdP risk categories. The most useful predictors were identified in the study: (1) the ability of a drug to induce “mild” (type A) or “severe” (all other) arrhythmia-like events at any concentration (predictor 1); (2) the extent of drug-induced repolarization prolongation at any concentration (predictor 4); and (3) the extent of drug-induced prolongation at the clinical C_{max} (predictor 7). We found it interesting that the ability of a drug to inhibit hiPSC-CMs' spontaneous beating or any of the other predictors did not further improve model prediction.

Despite the variations in the experimental protocols, including intended range of the tested EP platforms (5 different platforms; both MEA and VSO were used) and some unintended variations in cell batch, recording medium composition, and other parameters (Table S1), the results for all 28 drugs were fairly consistent across 10 sites. Table S6 shows the results of Pearson correlation analysis for drug-induced ddfPdc/APD90c change across 10 sites. Lower coefficient values for individual sites can be achieved by greater or lesser responses to drugs. Data from most of the sites were highly correlated (average Pearson coefficients of 78%–88%), while sites 2 and 4 had lower correlation coefficients (69% and 70%, respectively), potentially for different reasons because the Pearson coefficient between these 2 sites is low (37%). Differences observed for site 2 may be attributed to the differences in experimental protocols, because site 2 was the only test site that used the VSO platform instead of the MEA platforms and serum-free experimental medium instead of serum-containing medium. APD90 (VSO) and FPD (MEA) are equivalent measures, and site 2 showed appropriate APD changes and arrhythmia-like events in response to drugs, including dofetilide. However, the absence of serum in the assay media used by site 2 and the known potential of serum components (e.g., albumen) to modulate the bioavailability of some drugs in the serum-contain-

ing media of all of the other sites (Ando et al., 2017; Schocken et al., 2018) has the potential to explain the slightly lower average Pearson coefficient. In contrast, site 4 used experimental protocols that were largely consistent with other MEA sites, including the use of serum-containing media (Table S1), but overall correlation for that site was lower. Of note, Figure 4 shows that unlike all of the other sites, site 4 did not report large effects induced by dofetilide (no significant ddfPdc prolongation or drug-induced arrhythmia-like events). It will be critical to include positive drug controls (with known ion channel effects) on each plate to demonstrate suitable assay sensitivity based on the predominant mechanisms for affecting repolarization. It is important to recognize that the model demonstrates the ability of hiPSC-CMs across sites to detect delayed repolarization and predict TdP risk for 15 datasets for 28 drugs, but that an individual site may not be expected to detect the proarrhythmic risk for each drug.

Furthermore, despite different reprogramming and differentiation protocols used to manufacture the two hiPSC-CM lines used in the study, they were similar in predicting intermediate versus low or low versus high- or intermediate-risk drugs, which is the current unmet need. However, it is important to note that this study was limited to two hiPSC-CM lines and that other lines will require their own validation. Furthermore, the drugs can be potentially tested in gender-specific or even subject-specific hiPSC-CMs when feasible for the intended drug target population, but for this study we focused on a general assessment of a molecule, so the choice was made to use well-characterized, commercially developed hiPSC-CMs lines.

It is important to examine the outlier drugs that induced effects in hiPSC-CMs that are noticeably different from the other drugs in the same TdP risk category. Unlike other high-risk drugs and consistent with previous studies (Blinova et al., 2017), bepridil did not induce arrhythmias in hiPSC-CMs, even at 30-fold C_{max} (except for 1 of 15 datasets). Bepridil is a potent hERG blocker that also blocks L-type calcium and peak and late sodium currents at higher concentrations (Crumb et al., 2016). High expression levels of calcium ion channels in hiPSC-CMs as compared to primary ventricular tissue (Blinova et al., 2017) may have contributed to more attenuated cellular proarrhythmic effects of the drug as compared to other drugs in the high TdP risk category. It is also possible that the known propensity of bepridil to induce cardiac arrhythmia in the clinic is at least partly related to the ability of bepridil to affect hERG surface expression (Obejero-Paz et al., 2015). hERG trafficking effects of drugs were not assessed in this study because of the short duration exposures of hERG. Another outlier drug was low TdP risk ranolazine, which induced significant repolarization prolongation and arrhythmias in hiPSC-CMs, uncharacteristic for this risk category. While ranolazine blocks the hERG potassium channel and prolongs QTc, it is not associated with TdP risk because hERG block is balanced by significant late sodium current block (Johannesen et al., 2016). Lower expression levels of sodium channels and decreased late sodium current in hiPSC-CMs compared to primary human ventricular tissue (Blinova et al., 2017; Lemoine et al., 2017) may contribute to the apparent proarrhythmic effects of ranolazine in hiPSC-CMs. Similarly, lower densities of late sodium current in hiPSC-CMs may explain

mexiletine-induced arrhythmia-like events. Finally, another low-risk drug that induced arrhythmia-like events in hiPSC-CMs was metoprolol, a beta-blocker, the effects of which may not be appropriately modeled in uninnervated hiPSC-CMs monocultures.

The differences in cellular electrophysiology between native tissue and iPSC-cardiomyocytes has been well documented (Gibson et al., 2014; Hoekstra et al., 2012; Ma et al., 2011), with iPSC-cardiomyocytes possessing, in general, spontaneous activity, depolarized membrane potential (V_m), slower AP upstroke, and longer APD and FPD. It is unclear how these differences would translate into systematic or class-specific misclassifications, but it does speak to the need for specific calibration controls demonstrating assay sensitivity for sodium, calcium, and potassium currents blockade. Relative differences in ion channel and current levels in iPSC-CMs as compared to adult ventricular myocytes are likely the most important factors to improve the accurate prediction of TdP risk, especially with multichannel blocking drugs. With the development of new biotechnologies aimed at the development of more adult-like hiPSC-CMs (Wanjare and Huang, 2017; Yang et al., 2014), the predictivity of hiPSC-CMs assays is expected to further improve. Furthermore, other predictors of proarrhythmia risk may be added to the model based on their ability to differentiate drugs from the three categories examined here. For example, triangulation of the cardiac AP has previously been correlated with the ability to cause TdP experimentally (Hondeghem et al., 2001). Measurements of AP triangulation (corrected for AP duration) based on data from the VSO platform were correlated with TdP risk category (Figure S3) and may prove to be a useful additional descriptor in the future. As demonstrated here, hiPSC-CMs are an important new human *in vitro* model for the assessment of TdP risk, and their role in CiPA should be considered along with the recent advances in *in silico* modeling to predict proarrhythmic cardiotoxicity (Li et al., 2017; Passini et al., 2017). Computer models of TdP risk based on experimentally measured multichannel drug effects show high predictivity and would be an important primary step in proarrhythmic risk assessment, at least until iPSC-CMs become even better representations of adult human cardiac myocytes. The advantages of using readily available human-derived cardiomyocyte preparations need to be considered along with comparisons of the accuracy of cardiomyocytes (versus *ex vivo* or *in vivo* animal models) in predicting proarrhythmic risk when defining the optimal role of hiPSC-CMs in drug discovery.

Several experimental limitations of the study are worth noting. First, the free drug concentrations in hiPSC-CM experiments were not measured. As shown in Table S7, several drugs (e.g., disopyramide, azimilide, clarithromycin) were reported by multiple (but not all) sites as being poorly soluble in DMSO at the required concentrations. Thus, additional measures were taken, such as sonicating, warming at 37°C, or increasing the DMSO percentage. It has been shown (Schocken et al., 2018) that serum content in the cell culture medium used for drug dilution could affect drug solubility and availability. Although all of the sites followed the same nominal set of drug preparation instructions, measurements of drug concentrations in the experimental wells were not performed. Second, this study does not allow for

the measurement of the effect of drug metabolites, which can in some cases be more toxic than the parent drug (e.g., the metabolite of astemizole, desmethylastemizole [Vorperian et al., 1996]). Third, measuring the effects of hERG blockers on FPD can be challenging for some drugs because of the decrease in repolarization T-wave amplitude, in addition to the drug-induced FPD prolongation. Fourth, the effects of only short exposures (30 min) of drugs were assessed in this study, while some non-acute proarrhythmic effects (not the emphasis of CiPA) may require longer exposures to affect channel expression. Finally, this study was not statistically powered to investigate the effect of the electrophysiological device on the hiPSC-CM assay's predictivity of proarrhythmic drug potential. Table S8 contains information on the fraction of drugs correctly characterized into a TdP risk category from the data stratified by the EP platform. However, these data should be interpreted with caution because the study design does not allow for distinguishing the effects of the specific device from other effects introduced by the cell type or by the experimental site itself. Further studies are needed to investigate whether device choice would be an important consideration in improving preclinical TdP risk assessment by hiPSC-CM-based assays.

In summary, this study used statistical modeling to identify the most predictive endpoints of hiPSC-CMs assays in TdP risk assessment. Using only 3 endpoints, model 1 separated drugs into low-risk versus combined intermediate- and high-risk categories with an AUC value of 0.87 (87% predictivity) at the sample size we used, regardless of the type of hiPSC-CM used; model 2 separated drugs into 3 separate risk categories and showed a slightly lower AUC value of 0.82. Different thresholds with each model, which have associated sensitivity and specificity values, can be selected based on when the assay is being used in drug development. Because the goal of CiPA is to increase specificity and hiPSC-CMs will be used to check for missed or unanticipated effects, a threshold with a high specificity will be required. For example, a threshold of 0.8 in model 1 is associated with a specificity of 0.89 and a sensitivity of 0.63. If a drug is predicted to have a low risk in the *in silico* TdP risk metric, but is positive at this threshold, then it could be important to understand the reason for this discrepancy. If the drug has low proarrhythmic risk due to balanced multi-ion channel block, such as ranolazine with both hERG and late sodium current block, then this discrepancy would not be surprising. Such a result should not hinder progressing with clinical development, in which drug-induced QT prolongation and signs of balanced ion channel block (no J-Tpeak prolongation [Johannesen et al., 2016]) would still be assessed in first-in-human studies. Thus, it will be important to perform an integrated risk assessment, taking into account the different components of CiPA when implementing CiPA to improve specificity and provide more accurate predictions of clinical TdP risk, rather than solely focusing on hERG block and QT prolongation.

EXPERIMENTAL PROCEDURES

Study Sites and Platforms

Ten independent laboratories participated in the study, using any 1 of the 4 MEA platforms: Maestro (Axion BioSystems, abbreviated to "AXN" in this

paper), CardioECR (ACEA Biosciences, “ECR”), Multiwell (Multichannel Systems, “MCS”), and AlphaMED64 (Alpha MED Scientific, “AMD”), or the VSO platform: CelloPTIQ (Clyde Biosciences, “CLY”).

hiPSC-CMs

Two commercially available hiPSC-CM cell lines were used: iCell² (Cellular Dynamics International) and Cor.4U (NCardia). iCell² are normally cryopreserved at approximately day 30 of the differentiation (similar to [Ma et al. \[2011\]](#)); the production procedures for Cor.4U were not disclosed by the manufacturer. Sites were instructed to follow manufacturers’ recommendations for hiPSC-CM plating and maintenance, including cell culture plate coating, cell plating densities, and assay time window. Spontaneously beating, 100% confluent iPSC-CMs monolayers were used for drug testing. [Table S1](#) contains information on the specific cell lots, cell-handling details, and variations by experimental site.

Drug Dilution and Addition

Blinded drug powder was sent to all of the sites by the Chemotherapeutic Agents Repository of the National Cancer Institute, and stored at -20°C until the day of testing. Four concentrations of each drug were studied ([Table S2](#)). Four DMSO stocks for each drug concentration were prepared and either used on the same day or aliquoted and frozen. Concentrated (10 \times) testing solutions (50 \times for sequential dosing) for each concentration were prepared freshly on the day of testing by diluting DMSO stocks into experimental medium (serum-containing maintenance medium for MEA experiments and serum-free medium for VSO experiments; see the [hiPSC-CM](#) section for more details). Ten-fold dilution was achieved when drugs were added to the experimental well to attain the targeted concentration. For sequential dosing, DMSO concentrations were adjusted sequentially up to 0.1% at the highest concentration to achieve the targeted concentration of each drug. If insoluble compound was observed in DMSO or 10 \times testing stock solutions, then warming to 37°C and sonicating for 20 min was recommended. [Table S7](#) contains information on when these measures were taken to improve drug solubility.

MEA and VSO Recordings of Drug-Induced Effects in hiPSC-CMs

All MEA and VSO recordings were performed at 37°C . Single concentrations of each drug were tested in each experimental well by all of the sites, except site 10, where sequential additions were used. A 100% media change was performed in hiPSC-CMs 2–24 hr before baseline recordings. Media compositions used for MEA and VSO recordings are shown in [Table S1](#). Concentration effects of each drug were recorded in ≥ 5 replicates for 97% of the collected data. Experimental points collected with < 5 replicates are marked with a star in [Figures 3 and 4](#) and in [Supplemental Data S1, S2, and S3](#). Vehicle (0.1% DMSO) control wells were included on each plate. After baseline recording and drug addition, the plates were left to re-equilibrate for at least 30 min before recordings.

Data Analysis

Data Exclusion Criteria

The results were excluded from the analysis if baseline parameters for a specific well were outside the following pre-specified quality standards: (1) hiPSC-CMs baseline spontaneous beating rate had to be within the 20–90 beats per min range (i.e., 0.3–1.5 Hz), (2) the baseline beating rate had to be within 6 SDs calculated for the baseline beating rate on all of the wells on the given plate, (3) the coefficient of variation for the baseline beat period had to be $< 5\%$, and (4) the depolarization spike amplitude had to be > 0.3 mV (MEA recordings only). Based on these criteria, no more than 3% of wells were excluded from analysis.

Drug-Induced Changes in Repolarization and Arrhythmia-like Events

Fridericia’s formula ([Fridericia, 2003](#)) was used to correct hiPSC-CM action potential duration (APD) and field potential duration (FPD) dependence on beating rate (APDc, FPDc). While not thoroughly validated for hiPSC-CMs, this formula is widely used in these assays ([Blinova et al., 2017](#); [Clements and Thomas, 2014](#); [Ando et al., 2017](#); [Yamamoto et al., 2016](#)). Baseline- and vehicle-controlled ddFPDc and ddAPD90c at 90% repolarization were calculated by averaging all DMSO-treated wells on the plate for vehicle control. Drug-induced arrhythmia-like events were counted and classified in 1 of 4 categories (A–D), as illustrated in [Figure 1](#). The relation between hiPSC-CMs action potential and

field potential, including correspondence between different arrhythmia-like events recorded by MEA and VSO, have been described previously ([Asakura et al., 2015](#)). Combination of events (e.g., AB, AC, ABC, ABCD) was also observed and recorded. Several drugs inhibited spontaneous hiPSC-CMs contractions, leading to a quiescent state (Q). MEA and VSO instrument operators were blinded to the drug treatment during data collection and analysis.

Statistical Methods

Descriptive Analysis

The primary measurement was the averaged baseline- and vehicle-controlled, Fridericia rate-corrected ddFPDc/ddAPD90c at each concentration. Drug concentrations were treated as ordinal variables, in which the order mattered but not the difference between concentration values. ddFPDc/APD90c was not calculated and was designated as missing for the drug concentrations in which $\geq 50\%$ of the wells were arrhythmic after dosing. The concentration effects of each drug were recorded in ≥ 5 replicates for 97% of the collected data. Arrhythmia was a binary outcome and was designated as “Yes” if it occurred in at least one well at any concentration.

Modeling and Model Validation

Seven endpoints characterizing drug responses on hiPSC-CMs were used to build a linear regression model predicting the drug TdP risk category ([Table S4](#)). For the model development, drug-induced repolarization prolongation in hiPSC-CMs recorded with the MEA platform (ddFPDc) and the VSO platform (ddAPD90c) was considered equivalent. Cell type was treated as a fixed effect and experimental site was treated as a random effect in these models. The predictor selection procedure was based on model-fitting diagnostics of the AIC, the Bayesian information criterion (BIC), the AUC, and cluster analysis among continuous predictors. Model validation was achieved through cross-validation and bootstrapping. Statistical analysis was done using SAS (SAS Institute, Cary, NC) and R (RStudio, Boston, MA) software.

DATA AND SOFTWARE AVAILABILITY

Raw data from this study is available on the CiPA website: <http://cipaproject.org/data-resources/>

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, eight tables, and three data files and can be found with this article online at <https://doi.org/10.1016/j.celrep.2018.08.079>.

ACKNOWLEDGMENTS

The study was partly funded through a Food and Drug Administration Broad Agency Announcement (BAA) contract (FDABAA-15-00121) to the Health and Environmental Sciences Institute (HESI). This project has been funded in part with federal funds from the National Cancer Institute, NIH (contract HHSN26120080001E). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services or the Food and Drug Administration, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government. This research was supported in part by the Developmental Therapeutics Program in the Division of Cancer Treatment and Diagnosis of the National Cancer Institute. In addition, the study was partly funded through the Research on Regulatory Science of Pharmaceuticals and Medical Devices of the Japan Agency for Medical Research and Development (JP17mk0104027). We would like to acknowledge the individuals at the 10 volunteer sites who participated in this study, ACEA Biosciences (USA), Axion BioSystems (USA), Bristol-Myers Squibb (USA), Clyde Biosciences (UK), Eisai (Japan), Genentech (USA), Janssen Pharmaceutical (JNJ) (Belgium), LSI Medience (Japan), Merck (USA), and NMI (Germany); Cellular Dynamics International and Ncardia for providing hiPSC-CMs for the study; the Health and Environmental Sciences Institute for coordinating the study; and David Fraley, Wei Liu, Camilla Abdullaeva, and Suran Galappaththige for help with figures.

AUTHOR CONTRIBUTIONS

All authors contributed to manuscript preparation. In addition, D.M., G.S., M.B., H.R.L., U.K., H.Z., H.S., X.Z., K.S., T.O., and Y.K. coordinated hiPSC-CMs data collection at the 10 volunteer sites. J.P. and G.G. coordinated the multisite study. L.G. prepared the drug and dilution protocols for distribution to the sites. T.K.F. and R.K. defined the cell culture protocols for the study. G.G., D.G.S., and N.S. designed the study. G.G., Q.D., and K.B. consolidated and analyzed the data, performed the statistical modeling, and wrote the manuscript.

DECLARATION OF INTERESTS

The following co-authors are employed by or have declared interests in MEA or VSO platform manufacturer companies: G.S., Clyde Biosciences; D.M., Axion BioSystems; and X.Z., ACEA Biosciences. The following co-authors are employed by or have declared interests in stem cell manufacturer companies: T.K., Cellular Dynamics International; and R.K., Ncardia. G.G. is an employee of and a shareholder in the biopharmaceutical company AbbVie.

Received: January 5, 2018

Revised: April 30, 2018

Accepted: August 24, 2018

Published: September 25, 2018

REFERENCES

- Ando, H., Yoshinaga, T., Yamamoto, W., Asakura, K., Uda, T., Taniguchi, T., Ojima, A., Shinkyo, R., Kikuchi, K., Osada, T., et al. (2017). A new paradigm for drug-induced torsadogenic risk assessment using human iPSC cell-derived cardiomyocytes. *J. Pharmacol. Toxicol. Methods* **84**, 111–127.
- Asakura, K., Hayashi, S., Ojima, A., Taniguchi, T., Miyamoto, N., Nakamori, C., Nagasawa, C., Kitamura, T., Osada, T., Honda, Y., et al. (2015). Improvement of acquisition and analysis methods in multi-electrode array experiments with iPSC cell-derived cardiomyocytes. *J. Pharmacol. Toxicol. Methods* **75**, 17–26.
- Blinova, K., Stohman, J., Vicente, J., Chan, D., Johannesen, L., Hortigon-Vinagre, M.P., Zamora, V., Smith, G., Crumb, W.J., Pang, L., et al. (2017). Comprehensive translational assessment of human-induced pluripotent stem cell derived cardiomyocytes for evaluating drug-induced arrhythmias. *Toxicol. Sci.* **155**, 234–247.
- Clements, M., and Thomas, N. (2014). High-throughput multi-parameter profiling of electrophysiological drug effects in human embryonic stem cell derived cardiomyocytes using multi-electrode arrays. *Toxicol. Sci.* **140**, 445–461.
- Colatsky, T., Fermini, B., Gintant, G., Pierson, J.B., Sager, P., Sekino, Y., Strauss, D.G., and Stockbridge, N. (2016). The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative - update on progress. *J. Pharmacol. Toxicol. Methods* **81**, 15–20.
- Crumb, W.J., Jr., Vicente, J., Johannesen, L., and Strauss, D.G. (2016). An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel. *J. Pharmacol. Toxicol. Methods* **81**, 251–262.
- Fermini, B., Hancox, J.C., Abi-Gerges, N., Bridgland-Taylor, M., Chaudhary, K.W., Colatsky, T., Correll, K., Crumb, W., Damiano, B., Erdemli, G., et al. (2016). A new perspective in the field of cardiac safety testing through the Comprehensive In Vitro Proarrhythmia Assay paradigm. *J. Biomol. Screen.* **21**, 1–11.
- Food and Drug Administration (2017). FDA briefing document. Pharmaceutical Science and Clinical Pharmacology Advisory Committee meeting. March 15, 2017. Strategies, approaches, and challenges in model-informed drug development (MIDD). <https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/AdvisoryCommitteeForPharmaceuticalScienceandClinicalPharmacology/UCM544838.pdf>.
- Fridericia, L.S. (2003). The duration of systole in an electrocardiogram in normal humans and in patients with heart disease. 1920. *Ann. Noninvasive Electrocardiol.* **8**, 343–351.
- Gibson, J.K., Yue, Y., Bronson, J., Palmer, C., and Numann, R. (2014). Human stem cell-derived cardiomyocytes detect drug-mediated changes in action potentials and ion currents. *J. Pharmacol. Toxicol. Methods* **70**, 255–267.
- Hoekstra, M., Mummery, C.L., Wilde, A.A., Bezzina, C.R., and Verkerk, A.O. (2012). Induced pluripotent stem cell derived cardiomyocytes as models for cardiac arrhythmias. *Front. Physiol.* **3**, 346.
- Hondeghem, L.M., Carlsson, L., and Duker, G. (2001). Instability and triangulation of the action potential predict serious proarrhythmia, but action potential duration prolongation is antiarrhythmic. *Circulation* **103**, 2004–2013.
- Johannesen, L., Vicente, J., Mason, J.W., Erato, C., Sanabria, C., Waite-Labott, K., Hong, M., Lin, J., Guo, P., Mutlib, A., et al. (2016). Late sodium current block for drug-induced long QT syndrome: results from a prospective clinical trial. *Clin. Pharmacol. Ther.* **99**, 214–223.
- Jonsson, M.K., Vos, M.A., Mirams, G.R., Duker, G., Sartipy, P., de Boer, T.P., and van Veen, T.A. (2012). Application of human stem cell-derived cardiomyocytes in safety pharmacology requires caution beyond hERG. *J. Mol. Cell. Cardiol.* **52**, 998–1008.
- Kawakami, K., Nagatomo, T., Abe, H., Kikuchi, K., Takemasa, H., Anson, B.D., Delisle, B.P., January, C.T., and Nakashima, Y. (2006). Comparison of hERG channel blocking effects of various beta-blockers—implication for clinical strategy. *Br. J. Pharmacol.* **147**, 642–652.
- Lemoine, M.D., Mannhardt, I., Breckwoldt, K., Prondzynski, M., Flenner, F., Ulmer, B., Hirt, M.N., Neuber, C., Horváth, A., Kloth, B., et al. (2017). Human iPSC-derived cardiomyocytes cultured in 3D engineered heart tissue show physiological upstroke velocity and sodium current density. *Sci. Rep.* **7**, 5464.
- Li, Z., Dutta, S., Sheng, J., Tran, P.N., Wu, W., Chang, K., Mdluli, T., Strauss, D.G., and Colatsky, T. (2017). 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. Arrhythm. Electrophysiol.* **10**, e004628.
- Ma, J., Guo, L., Fiene, S.J., Anson, B.D., Thomson, J.A., Kamp, T.J., Kolaja, K.L., Swanson, B.J., and January, C.T. (2011). High purity human-induced pluripotent stem cell-derived cardiomyocytes: electrophysiological properties of action potentials and ionic currents. *Am. J. Physiol. Heart Circ. Physiol.* **301**, H2006–H2017.
- Millard, D., Dang, Q., Shi, H., Zhang, X., Strock, C., Kraushaar, U., Zeng, H., Levesque, P., Lu, H.R., Guillon, J.M., et al. (2018). Cross-site reliability of human induced pluripotent stem-cell derived cardiomyocyte based safety assays using microelectrode arrays: results from a blinded CiPA pilot study. *Toxicol. Sci.* **164**, 550–562.
- Obejero-Paz, C.A., Bruening-Wright, A., Kramer, J., Hawryluk, P., Tatalovic, M., Dittrich, H.C., and Brown, A.M. (2015). Quantitative profiling of the effects of vanoxerine on human cardiac ion channels and its application to cardiac risk. *Sci. Rep.* **5**, 17623.
- Passini, E., Britton, O.J., Lu, H.R., Rohrbacher, J., Hermans, A.N., Gallacher, D.J., Greig, R.J.H., Bueno-Orovio, A., and Rodríguez, B. (2017). Human *in silico* drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Front. Physiol.* **8**, 668.
- Sager, P.T., Gintant, G., Turner, J.R., Pettit, S., and Stockbridge, N. (2014). Re-channeling the cardiac proarrhythmia safety paradigm: a meeting report from the Cardiac Safety Research Consortium. *Am. Heart J.* **167**, 292–300.
- Sala, L., Bellin, M., and Mummery, C.L. (2017). Integrating cardiomyocytes from human pluripotent stem cells in safety pharmacology: has the time come? *Br. J. Pharmacol.* **174**, 3749–3765.
- Schocken, D., Stohman, J., Vicente, J., Chan, D., Patel, D., Matta, M.K., Patel, V., Brock, M., Millard, D., Ross, J., et al. (2018). Comparative analysis of media effects on human induced pluripotent stem cell-derived cardiomyocytes in proarrhythmia risk assessment. *J. Pharmacol. Toxicol. Methods* **90**, 39–47.
- Stockbridge, N., Morganroth, J., Shah, R.R., and Garnett, C. (2013). Dealing with global safety issues : was the response to QT-liability of non-cardiac drugs well coordinated? *Drug Saf.* **36**, 167–182.

Vorperian, V.R., Zhou, Z., Mohammad, S., Hoon, T.J., Studenik, C., and January, C.T. (1996). Torsade de pointes with an antihistamine metabolite: potassium channel blockade with desmethylastemizole. *J. Am. Coll. Cardiol.* *28*, 1556–1561.

Wanjare, M., and Huang, N.F. (2017). Regulation of the microenvironment for cardiac tissue engineering. *Regen. Med.* *12*, 187–201.

Yamamoto, W., Asakura, K., Ando, H., Taniguchi, T., Ojima, A., Uda, T., Osada, T., Hayashi, S., Kasai, C., Miyamoto, N., et al. (2016). Electrophysiological characteristics of human iPSC-derived cardiomyocytes for the assessment of drug-induced proarrhythmic potential. *PLoS One* *11*, e0167348.

Yang, X., Pabon, L., and Murry, C.E. (2014). Engineering adolescence: maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ. Res.* *114*, 511–523.

Cell Reports, Volume 24

Supplemental Information

**International Multisite Study of Human-Induced
Pluripotent Stem Cell-Derived Cardiomyocytes
for Drug Proarrhythmic Potential Assessment**

Ksenia Blinova, Qianyu Dang, Daniel Millard, Godfrey Smith, Jennifer Pierson, Liang Guo, Mathew Brock, Hua Rong Lu, Udo Kraushaar, Haoyu Zeng, Hong Shi, Xiaoyu Zhang, Kohei Sawada, Tomoharu Osada, Yasunari Kanda, Yuko Sekino, Li Pang, Tromondae K. Feaster, Ralf Kettenhofen, Norman Stockbridge, David G. Strauss, and Gary Gintant

Supplemental Table 1. Experimental conditions by site. Related to Experimental Procedures.

iCell² cardiomyocytes
Cor.4U cardiomyocytes

Parameter	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10
Cell lot #	11515	11515	11515	11515	11515	21573A	11515	11515	11515	11515
	CB434CL	CB434CL	CB434CL	CB434CL	CB434CL					
Platform	AXN	CLY	MCS	ECR	MCS	ECR	AXN	AXN	AXN	AMD
	AXN	CLY	AXN	ECR	MCS					
Plating density, 1000/well	50	25	36	30	36	50	72	50	70	30
	20	25	10	30	8					
Culture age on test day	6-10	7-8	7	7	7	7-8	7-8	7	7-8	6-8
	7-10	6-7	5	7	6					
Antibiotics used	none/G/P-S/A	G	G	G	none	none	P-S	none	P-S	P-S
	C	C	C	C	none					
Dosing scheme	single	single	single	single	single	single	single	single	single	seq
	single	single	single	single	single					
Recording media	CDI	SF1	CDI	CDI	CDI	CDI	CDI	CDI	CDI	CDI
	AXG	SF2	AXG	AXG	AXG					
Hours in recording media	19-25	12-18	4	18-24	4	18	4	4.5	4	2
	5-26	12-18	4	24	4					

Platforms: AXN=Maestro (Axion Biosystems), CLY=CellOPTIQ (Clyde Biosciences), ECR= CardioECR (ACEA Biosciences), AMD= AlphaMED64 (Alpha MED Scientific), MCS= MEA2100 (Multichannel Systems)

Antibiotics used in hiPSC-CM maintenance medium (not in the recording medium): G= gentamicin, P-S= penicillin/streptomycin, A= amphotericin B, C= ciprofloxacin. Dosing scheme: single (one drug concentration per well) or seq =sequential (increasing drug concentration in the same well) Recording media CDI iCell² maintenance medium (Cellular Dynamics), AXG=Cor.4U maintenance medium (Ncardia), SF1= DMEM (Gibco #11966-025) + 10mM galactose + 1mM sodium pyruvate, SF2=Serum-free Cor.4U medium (Ncardia). Plating areas (in mm²) were as following depending on a platform, AXN: 5 for iCell², 7 for Cor.4U; ECR: 19.6; CLY:32; AMD: 7.8; MCS: 0.45.

Supplemental Table 2. Drug concentrations. Related to Experimental Procedures.

Drug name	C _{max} , μM	Drug concentration, μM				Interval
		1	2	3	4	
High TdP risk						
Azimilide	0.07	0.01	0.10	1.00	10.00	log
Bepriidil	0.032	0.01	0.10	1.00	10.00	log
D,l Sotalol	15	0.1	1.0	10.0	100.0	log
Disopyramide	0.7	0.100	1.000	10.000	100	log
Dofetilide	0.002	0.0003	0.0010	0.0032	0.0100	1/2 log
Ibutilide	0.1	0.0001	0.0010	0.0100	0.1000	log
Quinidine	3	0.95	3.00	9.49	30.00	1/2 log
Vandetanib	0.3	0.01	0.10	1.00	10.00	log
Intermediate TdP risk						
Astemizole	0.0003	0.0001	0.001	0.01	0.1	log
Chlorpromazine	0.0345	0.09507	0.30043	0.94937	3	1/2 log
Cisapride	0.00258	0.00317	0.01001	0.03165	0.1	1/2 log
Clarithromycin	1.206	0.1	1	10	100	log
Clozapine	0.071	0.09507	0.30043	0.94937	3	1/2 log
Domperidone	0.02	0.003	0.03	0.3	3	log
Droperidol	0.016	0.03169	0.10014	0.31646	1	1/2 log
Ondansetron	0.372	0.03	0.30	3.00	30.00	log
Pimozide	0.00043	0.00095	0.003	0.00949	0.03	1/2 log
Risperidone	0.0018	0.00317	0.01001	0.03165	0.1	1/2 log
Terfenadine	0.000286	0.001	0.01	0.1	1	log
Low TdP risk						
Diltiazem	0.128	0.01	0.10	1.00	10.00	log
Loratadine	0.00045	0.00095	0.003	0.00949	0.03	1/2 log
Metoprolol	1.8	3.16912	10.0144	31.6456	100	1/2 log
Mexiletine	2.5	0.1	1	10	100	log
Nifedipine	0.0077	0.001	0.01	0.1	1	log
Nitrendipine	0.00302	0.00951	0.03004	0.09494	0.3	1/2 log
Ranolazine	1.948	0.1	1	10	100	log
Tamoxifen	0.021	0.09507	0.30043	0.94937	3	1/2 log
Verapamil	0.045	0.001	0.01	0.1	1	log

Supplemental Table 3. Site-induced variability. Related to Experimental Procedures.

Type of effects	Variability Source	Root of mean square error (ms)
Fixed	Test site	170
Fixed	Cell type	245
Fixed	Drug concentration	482
Random	Test site	36
Random	Other Errors	67

Contribution of test site as source of variability in ddFPDc/APD90c averaged for 28 drugs across 10 test sites. Test site was treated as either a fixed effect or random effect. In either case, test site ranked lower compared to cell type and drug concentration or the sum of other errors based on root of mean square errors (SR MSE).

Supplemental Table 4. Drug-induced electrophysiological effects in hiPSC-CMs used as drug risk category predictors in the model building. Related to Experimental Procedures.

#	Predictor Description	Predictor Type
1	Did drug induced arrhythmias occur at any concentration? (0=no arrhythmia, 1=type A arrhythmia, 2=any other arrhythmia type)	Categorical
2	Were drug-induced arrhythmias observed at any concentration in $\geq 40\%$ wells (typically in at least 2 out of 5 replicate wells) (0=no, 1=yes)	Binary
3	Repolarization prolongation (ms) at the first drug concentration with statistically significant ($p \leq 0.05$) prolongation or shortening	Continuous
4	Maximum repolarization change (ms) observed at any concentration	Continuous
5	Drug concentration (folds over C_{max}) at which the first statistically significant ($p \leq 0.05$) repolarization prolongation was first observed	Continuous
6	Drug concentration (folds over C_{max}) when drug-induced arrhythmias were first observed	Continuous
7	Drug-induced repolarization change (ms) at C_{max}	Continuous

Supplemental Table 5. Modeling results. Related to Figures 6-7.

Model 1: Logistic Regression
(Low risk vs. [intermediate or high])

Analysis of Maximum Likelihood Estimates				
Parameter	Outcome	Estimate	Standard Error	Wald Chi-Square
Intercept		-0.1311	0.158	0.6887
Predictor4		0.00687	0.00162	17.8953
Predictor7		0.0232	0.00743	9.7703
Predictor1	A	0.6583	0.4345	2.296
Predictor1	O	1.7944	0.4199	18.2643

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Predictor4	1.007	1.004	1.01
Predictor7	1.024	1.009	1.039
Predictor1 A vs N	1.932	0.824	4.526
Predictor1 O vs N	6.016	2.642	13.698

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	87.2	Somers' D	0.744
Percent Discordant	12.8	Gamma	0.744
Percent Tied	0	Tau-a	0.326
Pairs	38205	AUC	0.872

Model 2: Ordinal Regression
(Low risk vs. intermediate or low vs. high)

Analysis of Maximum Likelihood Estimates						
Parameter	Risk*	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	
Intercept	2	-2.1102	0.2173	94.3127	<.0001	
Intercept	1	-0.1211	0.1497	0.6548	0.4184	
Cell_Type	Cor.4U	0.2211	0.1094	4.0855	0.0433	
Predictor7		0.0338	0.00498	46.1116	<.0001	
Predictor4		0.00105	0.000740	1.9975	0.1576	
Predictor1	A	2	0.3865	0.3999	0.9340	0.3338
Predictor1	A	1	1.0551	0.4184	6.3601	0.0117
Predictor1	O	2	0.8737	0.3003	8.4639	0.0036
Predictor1	O	1	2.1732	0.4104	28.0457	<.0001
Intercept	2	-2.1102	0.2173	94.3127	<.0001	

Odds Ratio Estimates				
Effect	Risk*	Estimate	95% Wald Confidence Limits	
Cell_Type (AXG vs CDI)		1.556	1.013	2.389
Predictor7		1.034	1.024	1.045
Predictor4		1.001	1.000	1.002
Predictor 1 A vs N	2	1.472	0.672	3.223
Predictor 1 A vs N	1	2.872	1.265	6.522
Predictor 1 O vs N	2	2.396	1.330	4.316
Predictor 1 O vs N	1	8.787	3.931	19.640

Association of Predicted Probabilities and Observed Responses				
Percent Concordant		81.7	Somers' D	0.637
Percent Discordant		18.0	Gamma	0.638
Percent Tied		0.2	Tau-a	0.422
Pairs		57675	cAUC	0.819

* Risk: 1 for low vs. high, 2 for low vs. intermediate

Supplemental Table 6. Cross-site correlation in drug-induced ddFPDc/APD90c reported for 28 drugs in iCell². Related to Figures 2-4.

Test Site	1	2	3	4	5	6	7	8	9	10	Average
1	1	0.75	0.89	0.77	0.90	0.73	0.94	0.95	0.94	0.92	0.88
2	0.75	1	0.63	0.37	0.66	0.78	0.70	0.67	0.67	0.63	0.69
3	0.89	0.63	1	0.46	0.78	0.78	0.74	0.78	0.78	0.78	0.76
4	0.77	0.37	0.46	1	0.66	0.61	0.74	0.80	0.76	0.79	0.70
5	0.90	0.66	0.78	0.66	1	0.88	0.71	0.92	0.80	0.90	0.82
6	0.73	0.78	0.78	0.61	0.88	1	0.79	0.81	0.78	0.69	0.78
7	0.94	0.70	0.74	0.74	0.71	0.79	1	0.89	0.92	0.91	0.83
8	0.95	0.67	0.78	0.80	0.92	0.81	0.89	1	0.95	0.95	0.87
9	0.94	0.67	0.78	0.76	0.80	0.78	0.92	0.95	1	0.92	0.85
10	0.92	0.63	0.78	0.79	0.90	0.69	0.91	0.95	0.92	1	0.85

Pearson correlation coefficients between 10 sites are shown

Supplemental Table 7. Drugs solubility. Related to Experimental Procedures.

Drug name	Risk category	Number of sites (out of 10) that reported drug solubility issues addressed with sonication, 37°C incubation or doubled DMSO content (0.2%)
Disopyramide	H	6
Azimilide	H	5
Vandetanib	H	2
Dofetilide	H	1
Quinidine	H	1
Sotalol	H	0
Bepridil	H	0
Ibutilide	H	0
Clarithromycin	I	7
Ondansetron	I	5*
Risperidone	I	4
Domperidone	I	3
Astemizole	I	2
Cisapride	I	1
Clozapine	I	1
Pimozide	I	1
Terfenadine	I	0
Chlorpromazine	I	0
Droperidol	I	0
Verapamil	L	2
Diltiazem	L	1
Metoprolol	L	1
Ranolazine	L	1
Mexiletine	L	0
Loratadine	L	0
Nifedipine	L	0
Nitrendipine	L	0
Tamoxifen	L	0

* - one site noticed cloudiness after drug addition to the cells (no adjustments were done to the dilution protocol)

Supplemental Table 8. Model 1 prediction by platform. Related to Figures 6, 7.

Platform	Cell Type	TdP risk category	% of drugs categorized correctly		
			Cut-off=0.6	Cut-off=0.7	Cut-off=0.8
AMD	iCell ²	low	77.8	88.9	100.0
AMD	iCell ²	high or intermediate	84.2	73.7	63.2
AXN	Cor.4U	low	83.3	88.9	94.4
AXN	Cor.4U	high or intermediate	68.4	60.5	44.7
AXN	iCell ²	low	77.8	86.1	88.9
AXN	iCell ²	high or intermediate	84.2	76.3	64.5
CLY	Cor.4U	low	66.7	66.7	66.7
CLY	Cor.4U	high or intermediate	78.9	73.7	68.4
CLY	iCell ²	low	77.8	77.8	77.8
CLY	iCell ²	high or intermediate	89.5	89.5	73.7
ECR	Cor.4U	low	66.7	77.8	77.8
ECR	Cor.4U	high or intermediate	78.9	78.9	63.2
ECR	iCell ²	low	94.4	94.4	100.0
ECR	iCell ²	high or intermediate	73.7	63.2	60.5
MCS	Cor.4U	low	88.9	100.0	100.0
MCS	Cor.4U	high or intermediate	73.7	73.7	57.9
MCS	iCell ²	low	61.1	83.3	83.3
MCS	iCell ²	high or intermediate	94.7	86.8	73.7

Predicted TdP risk for low vs high and intermediate risk drugs combined by five EP platforms used in the study: ECR (ACEA CardioECR), AMD (AlphaMed), AXN (Axion), CLY (CelloPTIQ), and MCS (Multichannel Systems) at three cut-off risk values: 0.6, 0.7 and 0.8. The study was not statistically powered to compare the predictivity of each platform and variations in the table above can not be attributed solely to the platform performance, since platform-induced effects could not be separated from site-induced effects due to the study design.

Supplemental Figure 1: Model Prediction Example. Related to Figures 6, 7.

Example:

A compound was tested in iCell² cells using MEA platform. No compound-induced arrhythmias of any type were observed; averaged across replicates maximum observed compound-induced ddFPDc was 74 ms and averaged across replicates compound-induced FPDC prolongation was 0.34 ms at the expected clinical concentration (Cmax):

<i>Predictor 1</i>
0

<i>Predictor 4</i>
74

<i>Predictor 7</i>
0.34

<i>Cell Type</i>
0

Model 1:

Using Formula (1) and Supplemental Table 5 estimates for the predictors we find for P1:

$$\begin{aligned} \text{Logit}(P1) &= \text{Intercept} + (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) = \\ &= -0.1311 + 0.00687 \times 74 + 0.0232 \times 0.34 + 0 = 0.3852 \rightarrow P1 = 0.595 \end{aligned}$$

i.e. predicted probability of the compound to have high or intermediate TdP risk is 59.5% and probability to have a low TdP risk is 40.5%.

Model 2:

Using Formula (2) and Supplemental Table 5 estimates for the predictors we find for P2a (probability of high versus low TdP risk):

$$\begin{aligned} \text{Logit}(P2a) &= \text{Intercept} + (\text{Cell Type}) + (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) = \\ &= -2.1102 + 0.2211 \times 0 + 0.00105 \times 74 + 0.0338 \times 0.34 + 0 = -2.021 \rightarrow P2a = 0.117 \end{aligned}$$

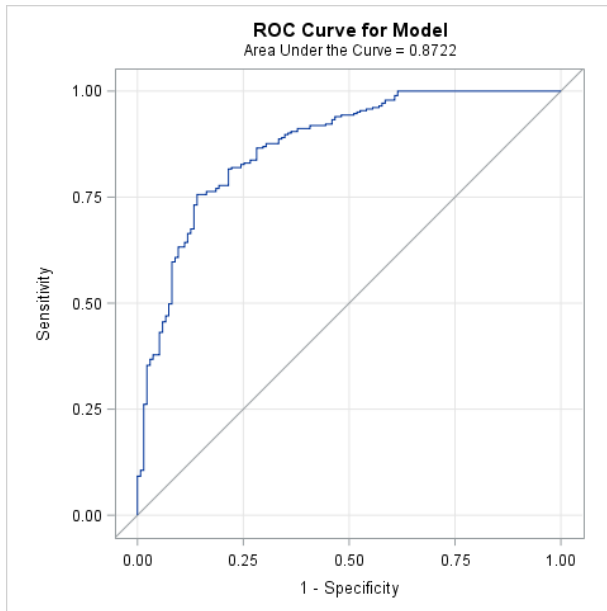
i.e. predicted probability of the compound to have high TdP risk is 11.7%.

Using Formula (3) and Supplemental Table 5 estimates for the predictors we find for P2b (probability of intermediate versus low TdP risk):

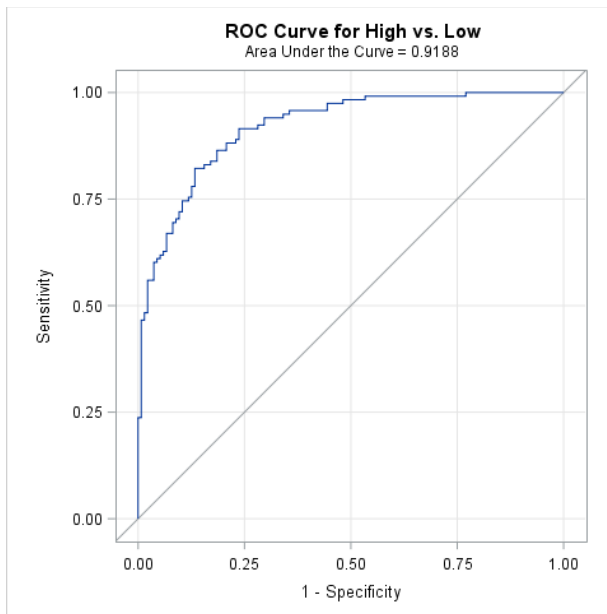
$$\begin{aligned} \text{Logit}(P2b) &= \text{Intercept} + (\text{Cell Type}) + (\text{Predictor}1) + (\text{Predictor}4) + (\text{Predictor}7) = \\ &= -0.1211 + 0.2211 \times 0 + 0.00105 \times 74 + 0.0338 \times 0.34 + 0 = 0.9686 \rightarrow P2a = 0.492 \end{aligned}$$

i.e. predicted probability of the compound to have intermediate TdP risk is 49.2% and the probability to have a low risk is 39.1%

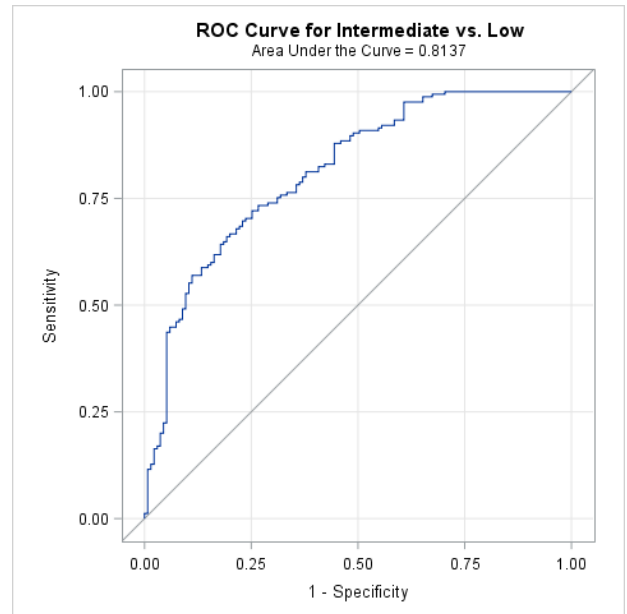
Supplementary Figure 2: ROC curves for Model 1 and 2. Related to Figures 6, 7.



ROC curve from logistic regression model (High & Intermediate vs. Low)



ROC curves from ordinal regression model (High vs. Low)



ROC curves from ordinal regression model (Intermediate vs. Low)

Supplemental Figure 3: Drug-induced changes in action potential morphology. Related to Figures 2-4.

