bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A deep learning algorithm to translate and classify cardiac electrophysiology: From iPSC-CMs to adult cardiac cells

Parya Aghasafari PhD[1], Pei-Chi Yang PhD[1], Divya C. Kernik PhD[2], Kauho Sakamoto PhD[4], Yasunari Kanda PhD[5], Junko Kurokawa PhD[4], Igor Vorobyov PhD[1,6] and Colleen E. Clancy PhD[1,3]

[1] Department of Physiology and Membrane Biology
University of California Davis, Davis, CA

[2] Washington University in St. Louis

[4]Department of Bio-Informational Pharmacology
School of Pharmaceutical Sciences
University of Shizuoka
Shizuoka, Japan

[5]Division of Pharmacology
National Institute of Health Sciences
Kanagawa, Japan

[6]Department of Pharmacology
University of California, Davis, CA

[3] Correspondence: Colleen E. Clancy, Ph.D.
University of California, Davis
Tupper Hall, RM 4303
Davis, CA 95616-8636
Email: ceclancy@ucdavis.edu

Phone: 530-754-0254

Word count:
Subject code: Artificial Intelligence, Machine Learning, Deep Learning, Computational Biology, Arrhythmias, Pharmacology.
Short title: Translation of cardiac action potential from immature to mature

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Exciting developments in both *in vitro* and *in silico* technologies have led to new ways to identify patient specific cardiac mechanisms. The development of induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) has been a critical *in vitro* advance in the study of patient-specific physiology, pathophysiology and response to drugs. However, the iPSC-CM methodology is limited by the low throughput and high variability of resulting electrophysiological measurements. Moreover, the iPSC-CMs generate immature action potentials, and it is not clear if observations in the iPSC-CM model system can be confidently interpreted to reflect impact in human adults. There has been no demonstrated method to allow reliable translation of results from the iPSC-CM to a mature adult cardiac response. Here, we demonstrate a new computational approach intended to address the current shortcomings of the iPSC-CM platform by developing and deploying a multitask network that was trained and tested using simulated data and then applied to experimental data. We showed that a deep learning network can be applied to classify cells into the drugged and drug free categories and can be used to predict the impact of electrophysiological perturbation across the continuum of aging from the immature iPSC-CM action potential to the adult ventricular myocyte action potential. We validated the output of the model with experimental data. The method can be applied broadly across a spectrum of aging, but also to translate data between species.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

The development of novel technologies has resulted in new ways to study cardiac function and rhythm disorders [1]. One such technology is the induced pluripotent stem cell-derived cardiomyocyte (iPSC-CMs) in vitro model system [2]. The iPSC-CM system constitutes a powerful *in vitro* tool for preclinical assessment of cardiac electrophysiological impact and drug safety liabilities in a human physiological context [3-8]. Moreover, because iPSC-CMs can be cultured from patient specific-cells, it is one of the new model systems for patient-based medicine.

While in vitro iPSC-CM utilization allows for observation of a variety of responses to drugs and other perturbations [9-12], there is still a major inherent limitation: The complex differentiation process to create iPSC-CMs results in a model of cardiac electrical behavior, which is relatively immature, resembling fetal cardiomyocytes. Hallmarks of the immature phenotype include spontaneous beating, immature calcium handling, presence of developmental currents, and significant differences in the relative contributions of repolarizing potassium currents compared to adult ventricular myocyte [13-15]. The profound differences between the immature iPSC-CM and the adult cardiac myocyte have led to persistent questions about the utility of the iPSC-CM action potential to predict relevant impact on adult human electrophysiology [16, 17]. In this study, we set out to take a first step to bridge the gap.

Here we describe a new way to connect data from patient derived iPSC-CMs to feed the development of computational models and to fuel the application of machine learning techniques to allow a way to predict, classify and translate changes observed in cardiac activity from the *in vitro* iPSC-CMs to predict their effects on adult cardiomyocytes (adult-CMs). The iPSC-CM and adult-CM action potential (AP) populations were paced with physiological noise with 0-50% I_{Kr} block to generate a robust simulated data set to train and test the deep learning algorithm. We then showed how it can be applied even to scarce experimental data, which was also used to validate the model.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Deep learning techniques are increasingly used to advance personalized medicine [18]. Long-short-term-memory (LSTM) based networks are capable of learning order dependence in sequence prediction problems [19] and have been widely used for cardiac monitoring purposes. They have been used to extract important biomarkers from raw ECG signals [20-22] and help clinicians to accurately detect common heart failure features in ECG screenings [20, 23-27]. LSTM classifiers have also been employed to automatically classify arrhythmias using ECG features [28-32]. In addition, modeling and simulation, and multi-task networks have been widely implemented to identify hERG blockers from sets of small molecules in drug discovery and screening [33-38].

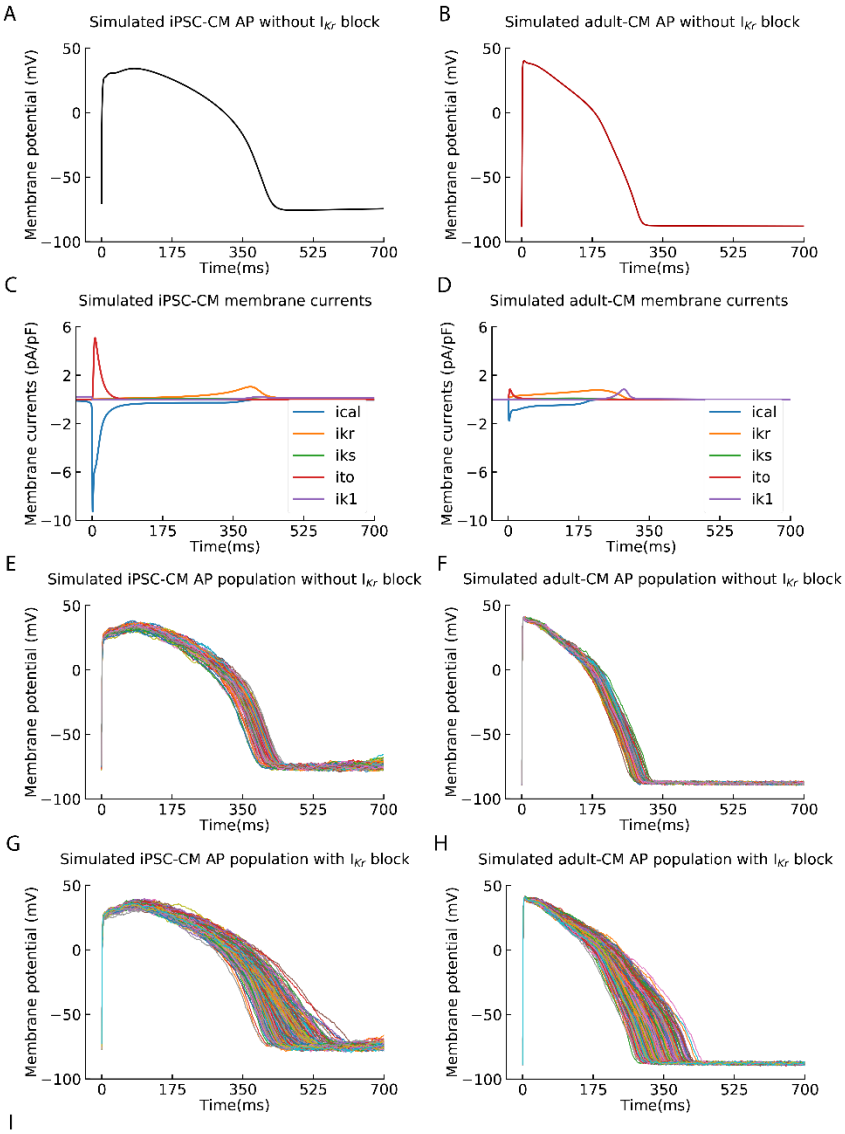
Here, we developed and applied a multitask network leveraging LSTM architecture to classify and translate observations from cardiac activity of in vitro iPSC-CMs to predict corresponding adult ventricular myocyte cardiac activity. Here, we show that developments in iPSC-CM experimental technology and developments in cardiac electrophysiological modeling and simulation of iPSC-CMs can be leveraged in a new machine learning model to interrogate disease and drug response in cardiac myocytes from immaturity to maturation.

Results

In this study, we set out to develop and apply a deep learning multitask network that would perform two distinct tasks: A) The first task is cellular level action potential classification to distinguish between drug free action potential waveforms and action potential waveforms following application of I_{Kr} block (drug-induced hERG channel block). B) The second goal is cellular level action potential translation to translate immature action potential waveforms to predict adult waveforms. We first simulated a population of 542 *in silico* iPSC-CMs and O'Hara-Rudy *in silico* human adult ventricular action potentials (adult-CMs). An average cell AP from the population is shown in Figure 1A for iPSC-CMs and Figure 1B for adult-CMs. The ionic currents underlying the *in silico* iPSC-CM APs and adult-CMs APs are markedly different as shown in panels C and D, respectively. In Figure 1E, the population of iPSC-CM APs is shown and the adult-CM AP population is in Figure 1F, generated by applying physiological noise as we have described previously [39-41]. In Figure 1 panels G and H, the impact of perturbation of the *in silico* iPSC-CM

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

APs and adult-CMs APs by 1-50% I_{Kr} block is shown. Panel I shows the parameter comparison for iPSC-CMs and adult-CMs APs without and with perturbation by 1-50% I_{Kr} block.



I

Measure	iPSC-CM (Δ)	adult-CM (Δ)
Without I_{Kr} block		
Upstroke Velocity (mV/ms)	13.2 – 25.8 (12.6)	247.3 – 271.3 (24)
MDP (mV)	-78.9 – -75.6 (3.3)	-86.5 - -89.5 (3)
APD_{30} (ms)	262 – 323 (61)	147.6 – 191.5 (43.9)
APD_{90} (ms)	376 – 439 (63)	253.3 – 307.9 (54.6)
With I_{Kr} block		
Upstroke Velocity (mV/ms)	13.8 – 25.1 (11.3)	247 – 275 (28)
MDP (mV)	-78.8 – -75.6 (3.2)	-88.7– -89 (0.3)
APD_{30} (ms)	283 – 369 (86)	168 – 226 (85)
APD_{90} (ms)	389 – 580 (191)	276 – 408 (132)

Figure 1. Cellular action potential (AP) and ionic currents for induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) and adult-CMs (O'Hara-Rudy human ventricular action

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

potentials). **A - B** Comparison of Cellular APs in the baseline model of iPSC-CMs and control case of adult-CMs at cycle length of 982 ms. **C**. Simulated ionic currents (I_{CaL}, I_{Kr}, I_{Ks}, I_{to}, I_{K1}) during baseline iPSC-CMs AP compared to the simulated currents profiles during control case of adult-CMs AP in **D**. **E**. iPSC-CMs APs of spontaneously beating cells (n = 304) were simulated after incorporating physiological noise. **F**. Adult-CMs APs were simulated with physiological noise currents at matching cycle lengths of iPSC-CMs in panel **E**. **G**. iPSC-CMs APs were simulated with perturbation by 1-50% I_{Kr} block after incorporating physiological noise. **H**. Adult-CMs APs were simulated with perturbation by 1-50% I_{Kr} block and physiological noise currents at matching cycle lengths of iPSC-CMs in panel G. **I**. Comparison between iPSC-CMs and adult-CMs for upstroke velocity, maximum diastolic potential (MDP) and action potential durations (APD).

We next developed a multitask deep learning network (Figure 2A), for the independent translation and classification tasks to be performed. The goal of the translation task (Figure 2B) is to use an immature cardiac action potential waveform dataset and convert (translate) these data to predict the resulting effect on mature cardiac action potential waveforms. Figure 2C shows the classification task that is intended to be used to classify action potential waveforms into drug free (iPSC-CMs APs with less than a determined threshold of (threshold%) I_{Kr} block) and drugged (iPSC-CMs APs with greater than or equal to a determined threshold of (threshold%) I_{Kr} block) categories. We combined these two tasks to demonstrate a positive relationship between planned tasks and empower the multitask network to predict mature cardiac action potential waveforms with better accuracy simultaneously. Both tasks are performed by a network comprising two long-short term memory (LSTM) layers (Figure 2D) followed by independent fully connected layers for each task (Figure 2E). Notably, the LSTM layers are shared for both tasks (please see methods for details on the LSTM layers). The features extracted by the LSTM layers are then used as input for the independent fully connected neural network layers with two hidden layers. The outputs from the network are both the mature cardiac action potential waveform, and the category of action potential waveform (drug free or drugged categories).

We used normalized and labeled *in silico* iPSC-CM APs and adult-CMs APs without (0% I_{Kr} block) (Figure 1 E-F) and with perturbation by 1-50% I_{Kr} block (Figure1 G-H) as input and output for training the multitask network (Figure 2A). First, we started training the network considering drug-free (0% block) and drugged (1-50% IKr) cases for classification task. Then, we tested if training the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

multitask network with cellular action potential waveforms that had been subject to various I_{Kr} blocking conditions (range 1-50% block) could improve the accuracy of the network for reconstructing adult-CM APs from iPSC-CM APs . The long short-term memory (LSTM) layer representation is shown as a schematic in Figure 2D. The computation module and gating mechanism for each training iteration is also depicted in Figure 2D.

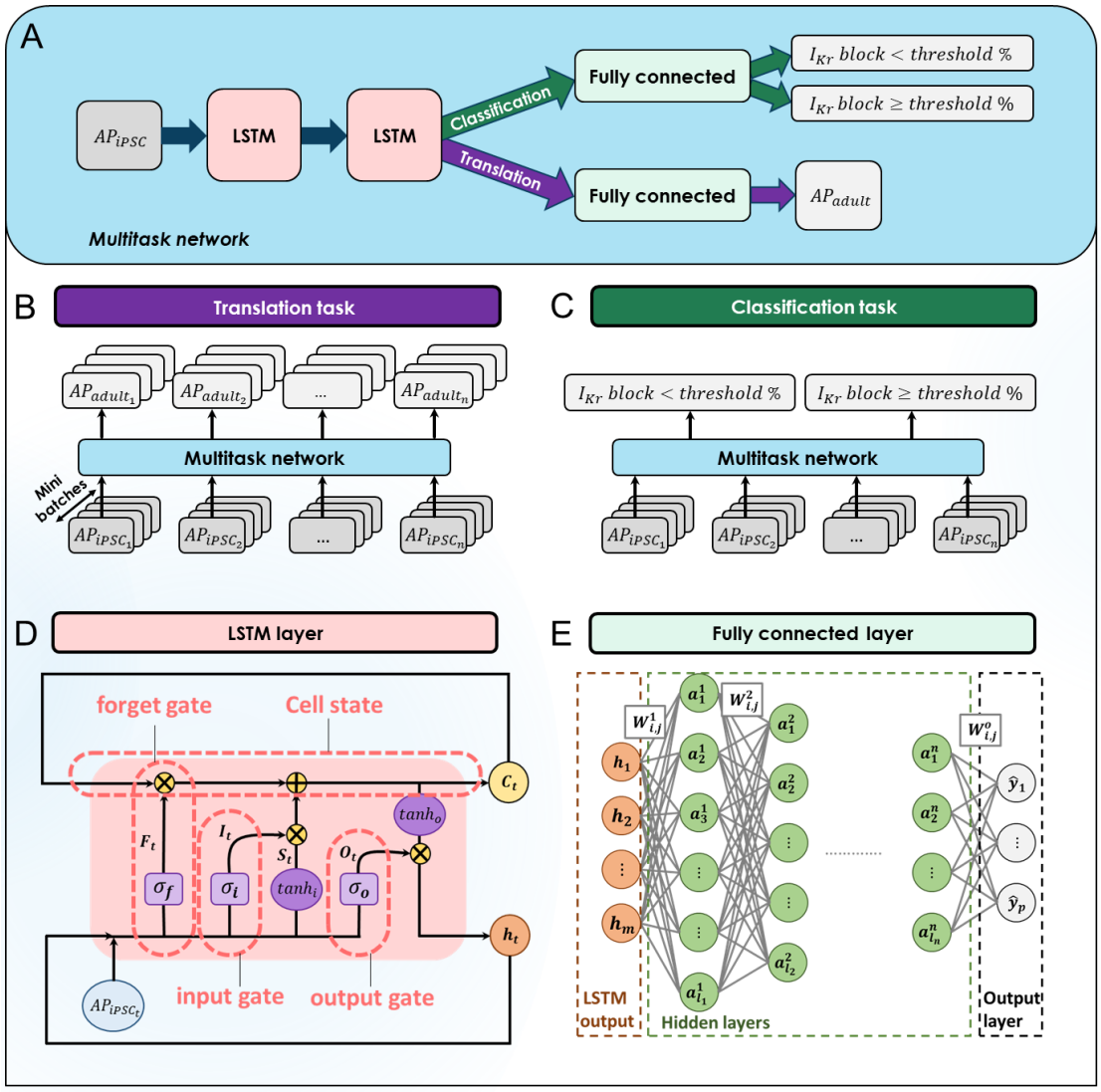


Figure 2. The multitask network architecture. **A**. The general overview of the multitask network presented in this study. **B**. The translation task to reconstruct adult-CMs APs from corresponding iPSC-CMs APs. **C**. The classification task is shown with I_{Kr} block threshold. **D**. The repeating module in the implemented LSTM layers in the presented multitask network. **E**. The architecture of the implemented fully connected layers in the presented multitask network.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We next explored whether the translation and classification tasks in this study are compatible with each other, which would allow them to be combined into a single multitask network. The benefits of combining the tasks are to save computation time and also potentially to improve each task performance. Therefore, we applied the multitask network to different scenarios in this study as shown in the workflow schematic in Figure 3. The data generated through the various scenarios are described in the subsequent figures in the study. We first trained the single task networks independently and compared the performance for the individual translation (purple, left) and classification (pink, second from left) tasks to the combined multitask network (yellow in middle). We applied statistical measures for the translation and classification tasks to evaluate task alignments. Next, we applied a forward and backward digital filter technique (light blue, second from right) to iPSC-CMs and adult-CMs APs traces and retrained the multitask network to explore the effect of noise filtering on the multitask network performance. In the last scenario, we investigated the importance of carrying out time dependency analysis by removing LSTM layers for both tasks (gray, right) and retrained the network with various threshold for the I_{Kr} block classification task (red block) to determine the optimal threshold for the best performance of the multitask network.

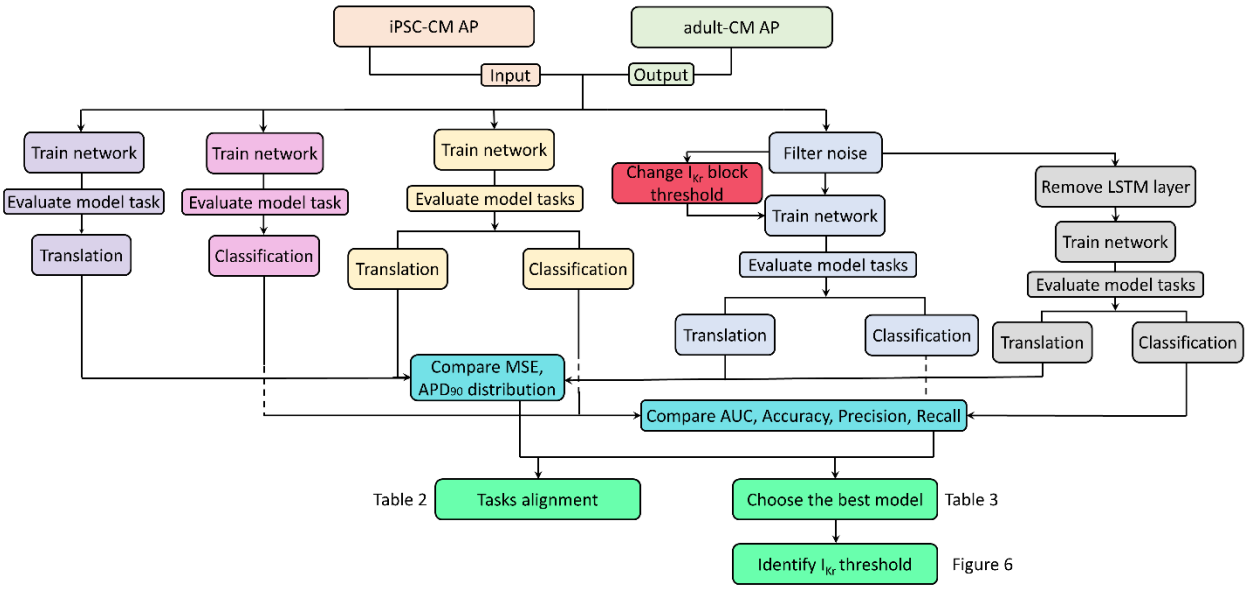


Figure 3. Work flow schematic describing different scenarios in this study: Exploring the task alignment by comparing the performance for the individual translation (purple, left) and classification (pink, second from left) tasks to the combined multitask network (yellow in middle);

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Inspecting the influence of applying noise-filtering techniques (light blue, second from right) to the iPSC-CMs and adult-CMs APs on the multitask network performance for the translation and classification tasks; Investigating the importance of considering the existing time dependency within simulated iPSC-CMs and adult-CMs APs for training the multitask network (gray, right); Tracking the performance of the multitask network for the translation and classification tasks with various threshold for the I_{Kr} block (red block) to identify the optimal threshold for classification task.

From the population of 542 computer model generated iPSC-CMs and adult-CM action potentials; which contained both drug free and drugged cases, we randomly selected 80% of the virtual cell waveforms containing both drug free and drugged cases to train the multitask network. The procedure was as follows: The multitask network utilized the iPSC-CM AP training data with physiological noise as an input to the network. The network was then optimized to best produce a translation of these data to match model generated adult-CM action potentials that were produced under the same pacing conditions and degree of I_{Kr} block (i.e. the immature and adult AP traces were generated through model simulations under the same pacing frequency conditions and extent of I_{Kr} reduction for both). These are shown as the "train" data set in Figure 4 A (iPSC-CMs) and B (adult-CMs training data in red). The optimized network was able to accurately translate the training data by generating a similar matched set of adult-CM action potentials as shown in blue in B (adult-CMs translated from iPSC-CMs by the network). The histogram in Figure 4C shows the good agreement between the model generated adult action potentials and iPSC-CM APs translated to adult-CM APs in terms of the frequency of virtual cells with similar APD.

The remaining 20% of virtual cell waveforms for both immature and mature APs were designated as the test set (data "unseen" to the network during training) and used to evaluate the performance of the multitask network as shown in Figure 4D (iPSC-CMs) and 4E for adult-CMs in red with superimposed translated data in blue. As indicated by the near superimposition of the histogram distribution in Figure 4 panel F, the translation from iPSC-CMs to adult-CMs AP across the range of pacing frequencies was successful. Since drug free and drugged data sets are mixed in the training and test set, we also plotted the simulated iPSC-CMs AP (Black) and adult-CMs AP (red) for drug free (Figure 4 G-H) and drugged AP waveforms (Figure 4J-K) separately to clarify the performance of the multitask network for translating iPSC-CMs to adult-CMs AP for categories

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

considered in the classification task. Superimposed translated adult-CMs AP waveforms are also illustrated in panel H and K in blue for drug free and drugged categories, respectively. Panel I and L illustrates the superimposition of the histogram distribution of APD_{90} values for drug free and drugged categories.

To evaluate the network performance for classification task we used area under receiver operating characteristic plot [42], recall, precision and F_1_score [43] as statistical measures (Table 1). We compared the values for discussed measures for both training and test set and observed that the network can categorize APs into drug-free and drugged waveforms with approximately 90% accuracy (Table 1).

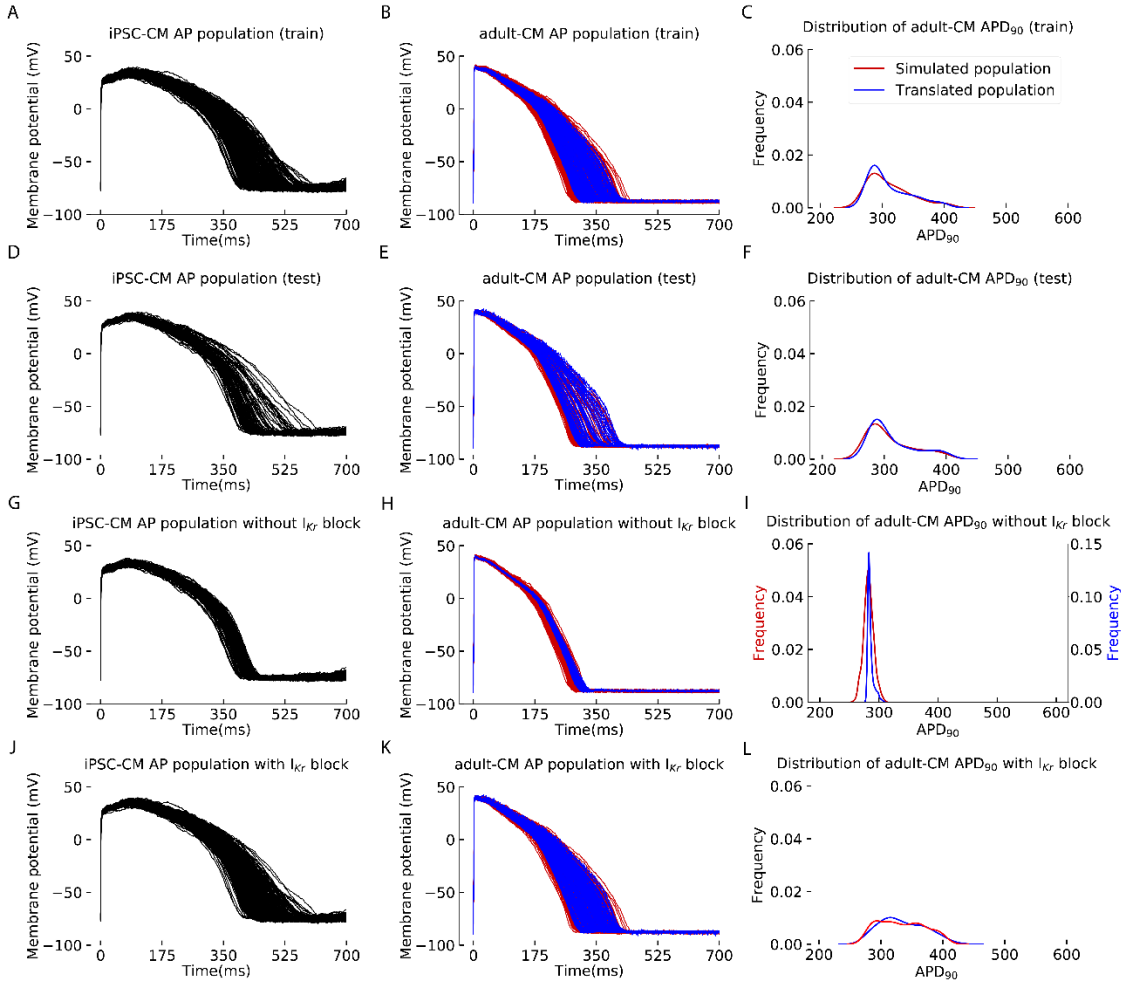


Figure 4. The performance of multitask network for translating iPSC-CMs APs into adult-CMs APs. **A.** The iPSC-CMs APs used for training the multitask network contained a variety of action potential

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

morphologies without and with I_{Kr} block (Train set). **B**. Comparison between simulated (red) and translated adult-CM APs (blue) in the training set. **C**. Comparison between the histogram distribution of APD_{90} values for simulated and translated adult-CM APs in the training set. **D**. Dedicated iPSC-CM APs for testing the performance of multitask network for the unseen AP data to the network (Test set) **E**. Comparison between simulated (red) and translated adult-CM APs (blue) in the test set. **F**. Comparison between histogram distribution of APD_{90} values for simulated and translated adult-CM APs in the test set. **G**. Population of iPSC-CM APs without I_{Kr} block. **H**. Comparison between simulated (red) and translated adult-CM APs (blue) without I_{Kr} block. **I**. Comparison between histogram distribution of APD_{90} values for simulated and translated adult-CM APs without I_{Kr} block. **J**. Population of iPSC-CMs APs with I_{Kr} block. **K**. Comparison between simulated (red) and translated adult-CM APs (blue) with I_{Kr} block. **L**. Comparison between histogram distribution of APD_{90} values for simulated and translated adult-CM APs with I_{Kr} block.

Table 1. Statistical measures for evaluating the performance of the multitask network for classifying AP traces into ones with and without 1-50% I_{Kr} block for both training and test sets.

Measures	AUROC	Recall	Precision	F_1-score
Training set performance	0.85	0.75	0.96	0.84
Test set performance	0.90	0.85	0.95	0.9

To evaluate the task alignment, we trained two single task networks for classification and translation tasks separately and compared the single task network performance with multitask network performance (Table 2). We compared the statistical measures for both translation and classification tasks. Interestingly, the multitask network actually performed better in drugged cases (Figure 4L) compared to drug-free cases (Figure 4I). As it can be seen in the presented measures, the multitask network has less mean-squared-error (MSE) and APD_{90} calculation error for translating the traces with 1-50% I_{Kr} block (Table 2). The implication is that addition of the classification task will improve the reliability of the network to translate iPSC-CMs APs into adult-CMs AP.

In contrast to the benefit of the multitask network to improving translation accuracy, no significant improvement was observed in statistical measures for classification task (Table 2). Notably, the statistical measures are close in value for the single task and multitask network which indicates

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

good task alignment in the multitask network. Therefore, combining the two tasks may indeed save computation time and also improve the translation task performance (Table2).

Table 2. Statistical measures for evaluating the performance of the multitask network for both translation and classification tasks to explore task alignment.

Translation				
Networks	MSE	R^2_score	APD_{90} error for block	
Single task (Translation)	0.0014	0.995	3.40%	
Single task (Classification)	-	-	-	
Multitask	0.0013	0.995	3.25%	
Classification				
Networks	AUROC	Recall	Precision	F_1-score
Single task (Translation)	-	-	-	-
Single task (Classification)	0.90	0.85	0.95	0.9
Multitask	0.90	0.85	0.95	0.9

Next, in order to demonstrate how the multitask network can be extended to apply even to noisy experimental data, we applied a digital filter forward-backward technique [44] to iPSC-CM and adult-CM AP traces without and with perturbation by 1-50% I_{Kr} block (Figure 5). Panel A shows simulated drug-free iPSC-CM with physiological noise in green and after applying a noise filtering technique in purple. To assess the phase distortion resulting from noise filtering for APD_{90} values, we then plotted a histogram distribution of APD_{90} values in Figure 5B, which shows near superimposition of the histogram distribution. This indicates that noise filtering does not change AP waveforms shape and only removes existing vertical noise. Panels C and D show the same process as described in panels A and B, but for the adult-CM APs (again, with physiological noise in green and after applying noise filtering techniques in purple in panel C, and histogram distributions of APD_{90} in panel D) .

In Figure 5E and F, we show the drugged AP waveforms with physiological noise in green, noise-filtered in purple and without physiological noise in pink for iPSC-CMs and corresponding histograms indicating the frequency of APD_{90} values, respectively. The adult-CM (Figure 5G) is shown for drugged AP waveforms with physiological noise in green, noise-filtered in purple and

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

without physiological noise in pink along with the overlaid APD_{90} histograms in Figure 5H. We also undertook retraining of the multitask network with noise-filtered traces to explore the effect of noise filtering on the network performance. We compared the statistical measures for both translation and classification tasks for the trained network based on AP waveforms with physiological noise and noise-filtered traces (Table 3) and observed that applying noise filtering techniques could slightly improve multitask network performance for both tasks.

Table 3. Statistical measures for evaluating the influence of applying noise-filtering techniques and considering the existing time dependency within simulated iPSC-CMs and adult-CMs APs on the multitask network performance.

Translation			
Networks	MSE	R^2_score	APD_{90} error with I_{kr} block
Multitask without noise-filtering	0.0013	0.995	3.25%
Multitask with noise-filtering	0.0013	0.995	3.17%
Multitask with noise-filtering and removing LSTM layers	0.0016	0.980	3.6%

Classification				
Networks	AUROC	Recall	Precision	F_1-score
Multitask without noise-filtering	0.90	0.85	0.95	0.90
Multitask with noise-filtering	0.91	0.87	0.95	0.91
Multitask with noise-filtering and removing LSTM layers	0.88	0.85	0.87	0.88

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

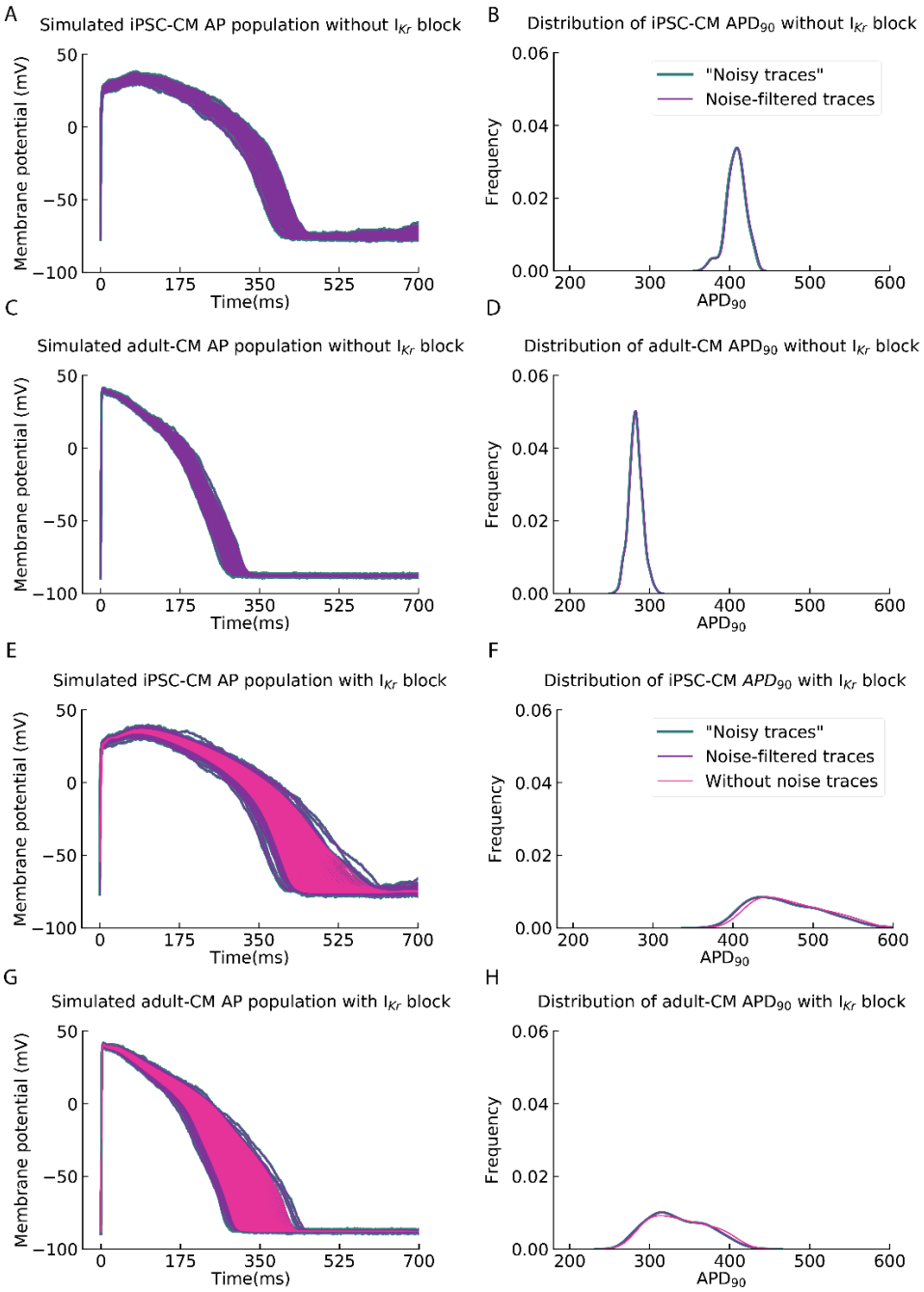


Figure 5. Applying a digital filter forward and backward technique to iPSC-CM and adult-CM APs without and with I_{Kr} block indicates zero phase distortion for APD_{90} values. **A**. Simulated drug-free iPSC-CMs APs with physiological noise in green and after applying the noise filtering technique in purple. **B**. Comparison between histogram distributions of APD_{90} values for noisy and noise-filtered iPSC-CMs APs. **C**. Simulated drug-free adult-CMs APs with physiological noise in green and after applying noise filtering technique in purple. **D**. Comparison between histogram distribution of APD_{90} values for noisy and noise-filtered adult-CMs APs. **E**. Simulated iPSC-CM APs with I_{Kr} block with physiological noise in green, after applying noise filtering technique on simulated traces in purple, and without physiological noise in pink. **F**. Comparison

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

between histogram distribution of APD_{90} values for noisy and noise-filtered versus iPSC-CM APs without physiological noise. **G**. Simulated adult-CM APs with I_{Kr} block with physiological noise in green, after applying noise filtering technique on simulated traces in purple, and without physiological noise in pink. **H**. Comparison between histogram distribution of APD_{90} values for noisy and noise-filtered versus simulated adult-CMs APs without physiological noise.

We next explored the importance of the existing time dependency within the simulated iPSC-CMs and adult-CMs APs on the multitask network performance. We did this by extracting the LSTM layers and compared the network performance with and without the LSTM layers for both translation and classification tasks (Table 3). As can be seen, removing LSTM layers worsened the network performance for both tasks. In summary, we determined that training the multitask network both with noise filtered traces and taking existing time dependence into consideration as the best approach to design and perform our desired translation and classification tasks.

To test whether training the multitask network with cellular action potential waveforms that had been subject to a variety of I_{Kr} blocking conditions (range 1-50% block) could improve the performance of the network, we retrained the multitask network with various I_{Kr} block percentages for the classification task and compared the performance of the multitask network to identify the optimal threshold for the classification task (Figure 6). For our test set, we recorded the MSE (Figure 6A) and the error in calculation of APD_{90} (Figure 6B) to investigate its influence on the translation task and monitored values of AUROC (Figure 6C) and precision (Figure 6D) for the classification task. As can be seen in Figure 6A-D, perturbation by 25% I_{Kr} block resulted in the best performance for both the translation and classification tasks when the selected threshold defined two classes: class 1 (iPSC-CMs APs with less than 25% I_{Kr} block) and class 2 (iPSC-CMs APs with greater than or equal to 25% I_{Kr} block). In summary, the multitask network with noised-filtered iPSC-CMs APs and adult-CMs APs with LSTM layers that defined a 25% I_{Kr} threshold for the classification task emerged as the best proposed network.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

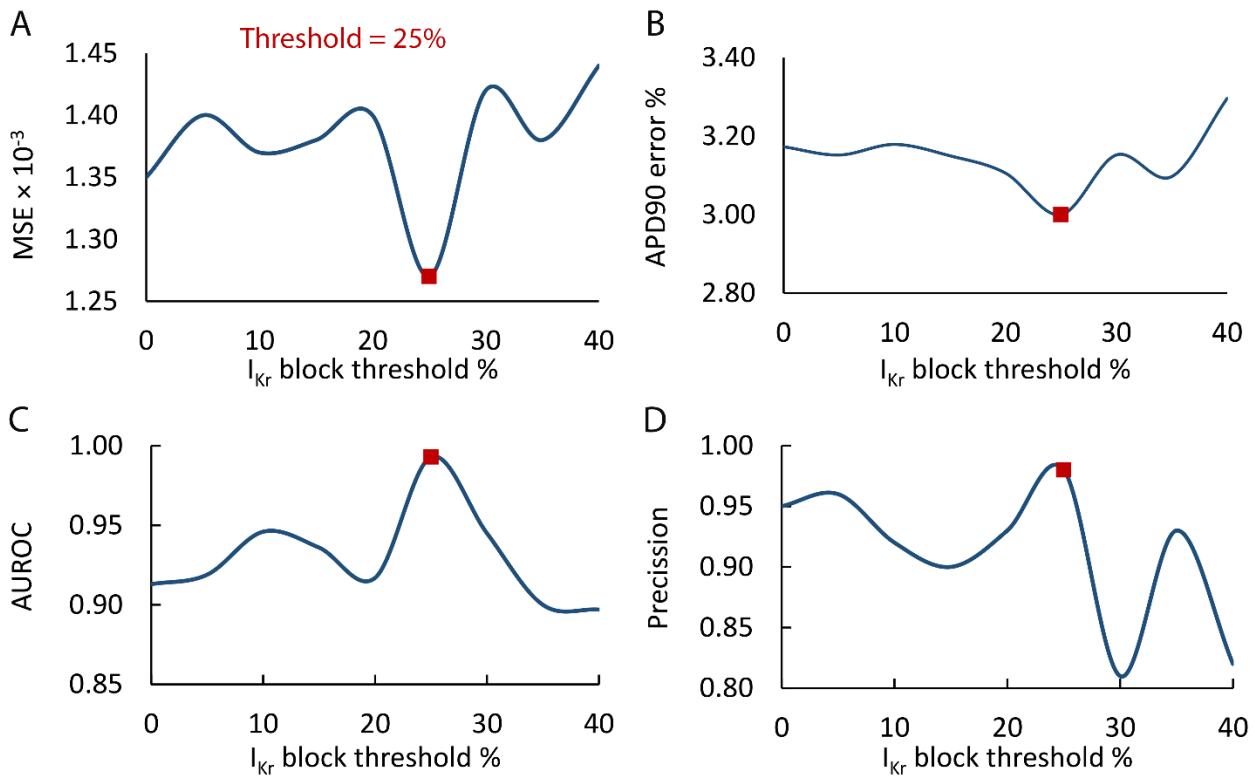


Figure 6. Tracking the performance of the multitask network for translation and classification tasks with different I_{Kr} block percentage thresholds to identify an optimal one. **A**. Calculated the mean-squared-error (MSE) between simulated and translated adult-CM AP waveforms by changing I_{Kr} block threshold for the classification task for test set. **B**. Calculated error between APD_{90} values for simulated and translated adult-CM APs for test set. **C**. Multitask network AUCROC measure for classifying iPSC-CM APs with I_{Kr} block into class 1 (iPSC-CM APs with less than threshold % I_{Kr} block) and class 2 (iPSC-CM APs with greater than or equal to threshold % I_{Kr} block). **D**. Multitask network precision measure for classifying iPSC-CM APs with I_{Kr} block into class 1 and 2.

We next set out to demonstrate the real-world utility of the multitask classification and translation network by applying the network to experimental data. We utilized experimental iPSC-CM APs from the Kurokawa lab (Figure 7A) and applied the translation task resulting in the predicted adult-CM APs as shown in Figure 7B. The translation notably resulted in a reduction in variability in APD in the adult translated cells, consistent with our simulated results and with previous experimental observations [16, 45]. In an additional validation of the deep learning network, we first simulated iPSC-CM APs with 50% block of I_{Kr}. We then used these simulation APs as an input for the multitask network and utilized the output from the translation task as a prediction of the effect of 50% block on adult-CMs. In Figure 7C, the translated drugged APD_{90} values are shown as turquoise asterisks

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

plotted with directly simulated O'Hara-Rudy adult-CMs APs with 50% I_{Kr} block (red curve) and additional validation by experimental 50% block of I_{Kr} by 1μM E-4031 (blue squares) [40]. These data strongly suggest that the effects of drug block in iPSC-CMs can successfully be translated to predict its effect on adult human APs.

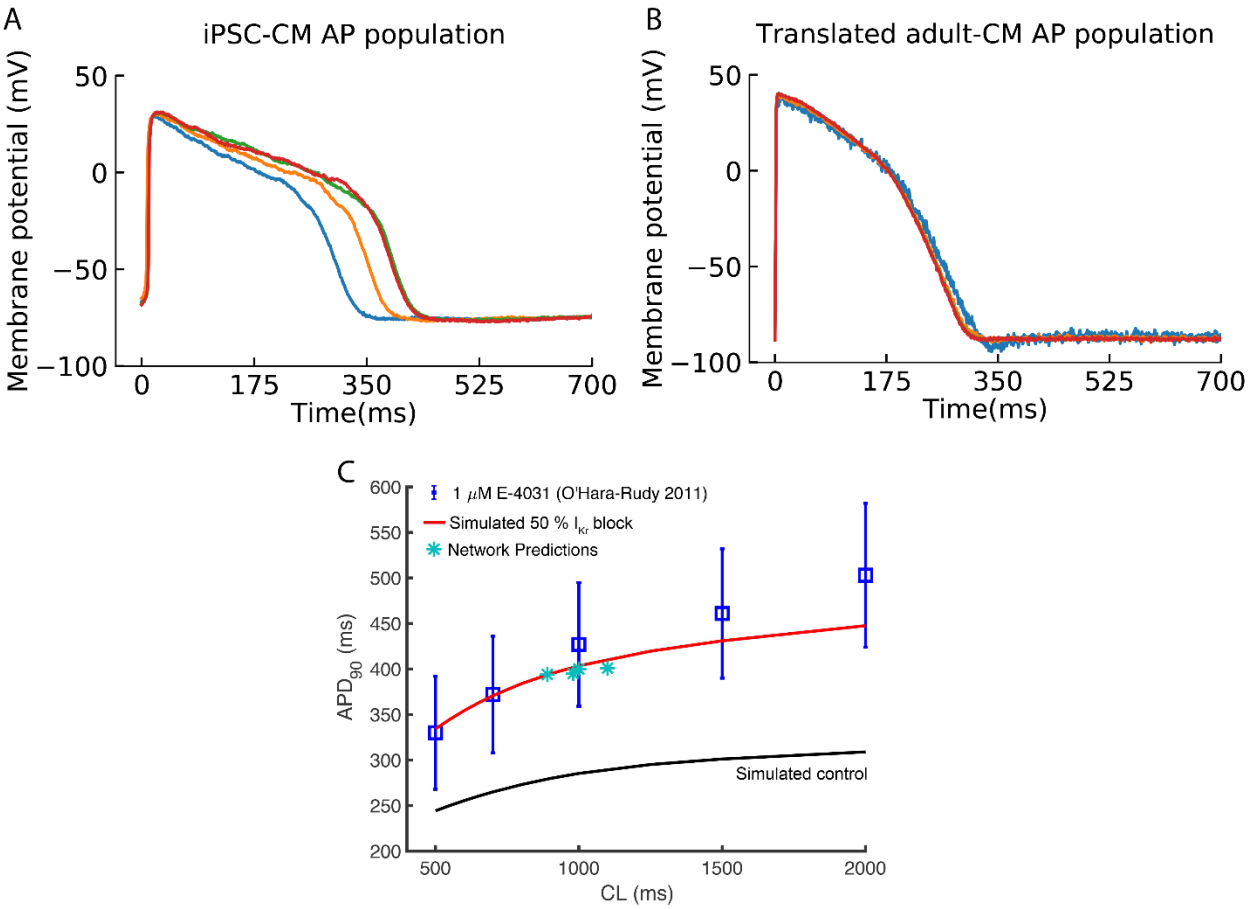


Figure 7. Translation of experimentally recorded iPSC-CMs APs into adult-CMs APs to validate the multitask network performance. **A**. Experimentally recorded iPSC-CMs APs from the Kurokawa lab. **B**. Translated adult-CMs APs from experimentally recorded iPSC-CMs APs using presented multitask network. **C**. Comparing translated adult-CMs AP APD_{90} values with 50% I_{Kr} block (turquoise asterisks) with previously published simulated (red curve for drugged and black for drug-free control) and experimental (blue squares) values in O'Hara-Rudy study [40] indicates successful model validation.

Methods

Generation of the *in silico* data for training and testing the deep learning model:

iPSC-CM and adult-CM baseline Action potentials with or without I_{Kr} block

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

iPSC-CMs baseline cells were paced from steady-state. The model formulation used for the baseline adult-CMs was the O'Hara-Rudy endocardial model cell [40]. The control adult-CMs was paced at cycle length of 982 ms to match the cycle length of the last beat of iPSC-CMs AP. The block case of adult-CMs (50% inhibition) was paced at cycle length of 1047 ms to match the cycle length of the last beat of iPSC-CMs AP with 50% I_{Kr} block.

iPSC-CMs and adult-CMs baseline simulations with physiological noise currents

Data set 1: Without I_{Kr} block: iPSC-CMs AP populations (n = 304) were generated after incorporating physiological noise. The adult-CMs were paced with noise for 100 beats after reaching steady state at matching cycle length of the last beat of iPSC-CMs AP populations.

Data set 2: With I_{Kr} block: The adult-CMs with noise were paced with 1-50% I_{Kr} block. The model was simulated at five varying beating rates for each percentage of block that match to the last beat of iPSC-CMs with 1-50% I_{Kr} block (n = 250). The numerical method used for updating the voltage was Forward Euler method [46].

Simulated physiological noise currents:

Simulated noise current was added to the last 100 paced beats in the simulated AP models, and simulated action potentials (APs) were recorded at the 2000th paced beat in single cells. This noise current was modeled using the equation from [39],

$$V_{t+\Delta t} = V_t - \frac{I(V_t)\Delta t}{C_m} + \xi n\sqrt{\Delta t} \quad (1)$$

Where n is N(0,1) is a random number from a Gaussian distribution, and Δt is the time step. ξ = 0.3 is the diffusion coefficient, which is the amplitude of noise. The noise current was generated and applied to membrane potential V_t throughout the last 100 beats of simulated time course.

Experimental iPSC-CMs:

Human iPSC-CMs (201B7, RIKEN BRC, Tsukuba, Japan) were cultured and subcultured on *SNL76/7* feeder cells as described in detail previously [47]. Cardiomyocyte differentiation was performed either by the EB method with slight modifications [47]. Commercially available iCell-cardiomyocytes (FUJIFILM Cellular Dynamics, Inc., Tokyo, Japan) were cultured according to the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

manual provided from the companies. Action potentials were recorded with the perforated configuration of the patch-clamp technique as described in detail previously [47]. Measurements were performed at *36 ± 1 °C* with the external solution composed of (in *mM*): *NaCl* (*135*), NaH_2PO_4 (*0.33*), *KCl* (*5.4*), $CaCl_2$ (*1.8*), $MgCl_2$ (*0.53*), glucose (*5.5*), *HEPES*, *pH 7.4*. To achieve patch perforation (*10-20 MΩ*; series resistances), amphotericin B (*0.3-0.6 µg/mL*) was added to the internal solution composed of (in *mM*): aspartic acid (*110*), *KCl* (30), $CaCl_2$ (*1*), adenosine-*5'*-triphosphate magnesium salt (*5*), creatine phosphate disodium salt (*5*), *HEPES* (*5*), *EGTA* (*11*), *pH 7.25*. In quiescent cardiomyocytes, action potentials were elicited by passing depolarizing current pulses (*2 ms* in duration) of suprathreshold intensity (*120 %* of the minimum input to elicit action potentials) with a frequency at *1 Hz* unless noted otherwise.

The multitask network architecture:

The proposed multitask network comprises two LSTM layers followed by independent fully connected layers (Figure 2A) for the classification and translation tasks. The LSTM layers memorizes the important time dependent information and then transfers the extracted information (features) into the subsequent fully connected layers to translate immature cardiac action potential waveforms into mature cardiac action potential waveforms (Figure 2B) and classify iPSC-CM APs into class 1 (iPSC-CM APs with less than the threshold % I_{Kr} block) and class 2 (iPSC-CM APs with greater than or equal to the threshold % I_{Kr} block) (Figure 2C).

Long-short term memory (LSTM) layers (Figure 2D):

We used LSTMs as the first two layers of the multi-task network to promote network learning for which data in a sequence is important to keep or to throw away. At each time step, the LSTM cell takes in three different pieces of information, the current input data (AP_{iPSC_t}), the short-term memory (hidden state) from the previous cell (h_{t-1}) and the long-term memory (cell state) (C_{t-1}). The LSTM cells contain internal mechanisms called gates. The gates are neural networks (with weights (w) and bias terms (b)) that regulate the flow of information at each time step before passing on the long-term and short-term information to the next cell [48]. These gates are called the input gate, the forget gate, and the output gate (Figure 2D).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The forget gate, as the name implies, determines which information from the long-term memory should be kept or discarded. This is done by multiplying the incoming long-term memory by a forget vector generated by the current input and incoming short-term memory. To obtain the forget vector, the short-term memory and current input are passed through a sigmoid function (σ) [49]. The output vector of sigmoid function is binary comprising 0s and 1s and is then multiplied by the long-term memory to choose which parts of the long-term memory to retain (Eq. 2).

$$F_t = \sigma_f\left(w_f AP_{iPSC_t} + w_f h_{t-1} + b_f\right) \quad (2)$$

The input gate decides what new information will be stored in the long-term memory. It considers the current input and the short-term memory from the previous time step and transforms the values to be between 0 (unimportant) and 1 (important) using a sigmoid activation function (Eq. 3). The second layer in input gate takes the short-term memory and current input and passes it through a hyperbolic tangent (tanh) activation function to regulate the network (Eq. 4).

$$I_t = \sigma_i\left(W_i AP_{iPSC_t} + W_i h_{t-1} + b_i\right) \quad (3)$$

$$S_t = tanh_i\left(w_s AP_{iPSC_t} + w_s h_{t-1} + b_s\right) \quad (4)$$

The outputs from the forget and input gates then undergo a pointwise addition to give a new version of the long-term memory (Eq . 5), which is then passed on to the next cell.

$$C_t = F_t * C_{t-1} + I_t * S_t \quad (5)$$

Finally, the output gate utilizes current input and previous short-term memory and passes them into a sigmoid function (Eq. 6). Then the new computed long-term memory passes through a tanh activation function and the outputs from these two processes are multiplied to produce the new short-term memory (Eq. 7).

$$O_t = \sigma_o(w_o x_t + w_o h_{t-1} + b_o) \quad (6)$$

$$h_t = O_t * tanh_o(C_t) \quad (7)$$

The short-term and long-term memory produced by these gates is carried over to the next cell for the process to be repeated. The output of each time step is obtained from the short-term memory,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

also known as the hidden state, and is subsequently passed into fully connected layers to perform the translation and classification tasks.

Fully connected layers (Figure 2E):

The fully connected neural network layers contains input, hidden and output layers (Figure 2E) which may have various numbers of neurons. Every neuron in a layer is connected to neurons in another layer [50]. Fully connected layers receive the output of LSTM layers as input. The fully connected layers calculate a weighted sum of LSTM outputs and add a bias term to the outputs. These data are then passed to an activation function (*f*) to define the output for each node (Eq. 8,9) [51].

$$a_j^n = f(Z_j^n) \tag{8}$$

$$Z_j^n = W_{i,j}^n * a_j^{n-1} + b^n \tag{9}$$

where a^0 is the input data and a^{n+1} is output data $\hat{y} \epsilon \left\{y_{t_i}, y_{c_i}\right\}$ where y_{t_i} and y_{c_i} are the outputs for translation and classification tasks, respectively. We first assign random values to all network parameters (θ_t; nodes weight $(W_{i,j})$, bias term (b) and network hyperparameters and select the best network infrastructure; number of hidden layers, number of neurons and activation functions for each hidden layer. Next, we estimate the network errors using mean-squared-error (Eq. 10) and cross-entropy loss functions (Eq. 11) to map the translation and classification tasks [52, 53], respectively.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \left\|y_{t_i} - \hat{y}_{t_i}\right\|^2 \tag{10}$$

$$CrossEntropy = -\left(y_{c_i}\log(\hat{y}_{c_i}) + (1 - y_{c_i})\log\left(1 - \hat{y}_{c_i}\right)\right) \tag{11}$$

where *n* is the total number of input samples and y_{t_i} and $\hat{y}_{t_i}$ are the simulated and translated adult-CM APs (the network output for translation task). The y_{c_i} is binary indicator of class labels for iPSC-CM APs (0 for iPSC-CM APs with less than the threshold % I_{Kr} block or 1 for iPSC-CM APs with greater than or equal to the threshold % I_{Kr} block) and $\hat{y}_{c_i}$ is predicted probability of APs being classified into the discussed classes. We used sum of both loss functions (Eq. 12) to calculate the overall network error (*J*) for both translation and classification tasks during the network training process. We updated network parameters (θ_{t+1}) using adaptive momentum estimation (ADAM)

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

optimization algorithm [54] based on the average gradient of overall loss function with respect to the network parameters for 128 randomly selected simulated AP traces (mini-batch=128) at each training iteration (Eqs. 13-15).

$$J(\theta_{t,i}) = CrossEntropy_{Classification}(\theta_{t,i}) + \mathrm{MSE}_{Translation}(\theta_{t,i}) \quad (12)$$

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\alpha . \hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad , \quad \theta_{t,i} \in \{W_{i,j}^n, b_j{}^n\} \quad (13)$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \text{where} \quad m_t = (1-\beta_1)\nabla J(\theta_{t,i}) + \beta_1 m_{t-1} \quad (14)$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t}, \quad \text{where} \quad v_t = (1-\beta_2)\left(\nabla J(\theta_{t,i})\right)^2 + \beta_2 v_{t-1} \quad (15)$$

We used rectified linear unit (ReLu) activation function for hidden layers and dropout regularization [55] with probability of eliminating any hidden units in fully connected layers equal to 0.25 for updating model parameters to find global minimum of loss function based on a predefined learning rate (α) , first and second momentum terms (β_1, β_2), and a small term preventing division by zero (ϵ).

We first started training the network considering drug-free (0% block) and drugged (1-50% I_{Kr}) cases for classification task. Then, we tested a hypothesis whether training the multitask network with classifying cellular action potential waveforms that had been subject to various I_{Kr} blocking conditions (range 1-50% block) could improve the performance of the network. We used MSE (Eq. 10), R^2_score (Eqs. 16-17 below) and the histogram distribution of APD_{90} as statistical measures to evaluate the performance of network for translation task and area under receiver operating characteristic curve (AUROC), recall, precision and F_1-score to measure capability of network for classification task as described below.

We applied a forward and backward digital filter technique[44] into normalized and labeled simulated iPSC-CM APs and adult-CM APs without and with perturbation by 1-50% I_{Kr} block and utilized them as inputs and outputs in the network. The network architecture is implemented using Pytorch platform[56]. 80% of the simulated APs were considered for training the network and the remaining were used to test the performance of the network for an unseen dataset during

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

training. The network codes have been made publicly available at Github. (https://github.com/ClancyLabUCD/Multitask_network)

Statistical measures

As we discussed, we used MSE and cross-entropy as criteria for performance evaluation of translation and classification tasks. In addition to MSE, we computed R^2_score [57] (Eqs. 16,17) to measure how close the translated adult-CM AP $(\hat{y}_{t_i})$ are to the simulated adult-CM AP (y_{t_i}). We compared the histogram distribution of simulated and translated adult-CM APD_{90} values to investigate the ability of network to predict them accurately.

$$\bar{y}_t = \frac{1}{n}\sum_{i=1}^{n} y_{t_i} \tag{16}$$

$$R^2 = \frac{\sum_i(\hat{y}_{t_i} - \bar{y}_t)}{\sum_i\left(y_{t_i} - \bar{y}_t\right)} \tag{17}$$

We used area under the receiver operating curve (AUROC) to measure the capability of model in distinguishing between classes [42]. Receiver operating curve (ROC) is a plot of the false positive rate (FPR, the probability that the network classifies iPSC-CM AP with less than 25% I_{Kr} block into greater than or equal to 25% I_{Kr} block) (Eq. 18) versus the true positive rate (TPR) or recall (Eq. 19). AUROC close to 1 represents an excellent model, which has good measure of separability, while a poor model has AUROC near 0, which means that it has poor separability.

prediction [58]. In addition, the confusion matrix elements were used to measure statistical metrics (recall, precision and F1-score) to describe the performance of a classification model [13], where recall (Eq. 19) is the proportion of actual positives that are correctly identified as such. Precision measures the proportion of correct positive identifications (Eq. 20) and F1-score is the harmonic mean of precision and recall (Eq. 21).

$$\mathrm{FPR} = \frac{FP}{FP+TN} \tag{18}$$

$$\mathrm{Recall} = \frac{TP}{TP+FN} \tag{19}$$

$$Precision = \frac{TP}{TP+FP} \tag{20}$$

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{21}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

In this study, we have developed a data-driven algorithm intended to address well known shortcomings in the induced pluripotent stem cell-derived cardiomyocyte (iPSC-CMs) platform. A known concern with iPSC-CMs is that the data collection results in measurements from immature action potentials, and it is unclear if these data reliably indicate impact in the adult cardiac environment. Here, we set out to demonstrate a new deep learning algorithm to allow reliable translation of results from the iPSC-CM to a mature adult cardiac response. The translation task yielded superb results with 0.0012 MSE, 0.99 R^2_score and less than 3% calculation error in predicting adult APD_{90} values from iPSC-CMs inputs.

The multi-task network we present here notably conferred additional benefit over considering the translation and classification tasks separately. For example, we noted that adding the classification task to distinguish data from action potentials with and without the application of an I_{Kr} blocking drug could improve the performance of the network. Adding the I_{Kr} blocking threshold to the translation task resulted in up to 7% decrease in MSE value. To pick the best threshold, we tested a range of thresholds for classification and retrained the network with each of them. The threshold defined as 25% I_{Kr} block led to the highest accuracy for both the translation and classification tasks. The high values for the classification task statistical measures including AUROC and precision equal to 0.993 and 0.98, respectively, also indicate the credibility of the multitask network to sort iPSC-CMs APs into groups with greater than or equal to 25% I_{Kr} block from iPSC-CMs APs with less than 25% I_{Kr} block.

Importantly, the multitask network presented here performed well even in the setting of the noted variability in measurements from iPSC-CMs. We utilized a modeling and simulation approach from our recent study [41] to generate a population of IPSC-CM action potentials that incorporate variability comparable to that in experimental measurements. Utilizing simulated data presented a unique opportunity: We were able to generate large amounts of data that were used both to train and optimize the network and then to test the network with specifically designated distinct

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

simulated data sets. ***Utilizing simulated data to train a deep learning network may constitute a more widely applicable approach that could be used to train variety of networks to perform multiple functions where access to comparable experimental data is not feasible***.

Following the optimization and demonstration of the network as an accurate tool for both translating and classifying data, we then used the same network to translate experimentally obtained data. We showed that the proposed network can effectively take experimental data as an input from immature iPSC-CMs and translate those data to produce adult action potential waveforms. It is notable that the variation observed in the adult-CM AP duration is smaller compared to iPSC-CM AP (Figure 7A-B). This has been observed both experimentally [16, 45] and in our simulated cell environment [41, 59] . Although the simulated iPSC-CM has a large initial calcium current (Figure 1B) compared to the simulated adult-CM (Figure 1C), the amplitude of currents flowing through adult-CM action potential plateau is notably larger. The immature iPSC-CM cells have low conductance during the AP plateau rendering it comparably higher resistance. For this reason, small perturbations to the iPSC-CMs have a larger impact on the resulting AP duration than observed in adult cells [60]. We also used simulated iPSC-CMs subject to 50% block of I_{Kr}. We translated those data to adult-CM APs and then compared the previously reported impact on adult-CM APs to 50% I_{Kr} block from experiments and noted excellent agreement thereby providing validation using experimental data from adult human cells.

In this study, we show that a deep learning network can be applied to classify cells into the drugged and drug free categories and can be used to predict the impact of electrophysiological perturbation across the continuum of aging from the immature iPSC-CM action potential to the adult ventricular myocyte action potential. We translated experimental immature APs into mature APs using the proposed network and validated the output of some key model simulations with experimental data. The multitask network in this study was used for translation of iPSC-CMs to adult APs, but could be readily extended and applied to translate data across species and classify data from a variety of systems. Also, another extension of the technology presented here is to predict the impact of naturally occurring mutations and other genetic anomalies [61].

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Shaheen, N., et al., *Human induced pluripotent stem cell-derived cardiac cell sheets expressing genetically encoded voltage indicator for pharmacological and arrhythmia studies.* Stem cell reports, 2018. **10**(6): p. 1879-1894.
2. Leyton-Mange, J.S., et al., *Rapid cellular phenotyping of human pluripotent stem cell-derived cardiomyocytes using a genetically encoded fluorescent voltage sensor.* Stem cell reports, 2014. **2**(2): p. 163-170.
3. Sun, N., et al., *Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy.* Sci Transl Med, 2012. **4**(130): p. 130ra47.
4. Lan, F., et al., *Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells.* Cell Stem Cell, 2013. **12**(1): p. 101-13.
5. Burridge, P.W., et al., *Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity.* Nat Med, 2016. **22**(5): p. 547-56.
6. Doss, M.X. and A. Sachinidis, *Current challenges of iPSC-based disease modeling and therapeutic implications.* Cells, 2019. **8**(5): p. 403.
7. Collins, T.A., M.G. Rolf, and A. Pointon, *Current and future approaches to nonclinical cardiovascular safety assessment.* Drug Discovery Today, 2020.
8. Wu, J.C., et al., *Towards precision medicine with human iPSCs for cardiac channelopathies.* Circulation research, 2019. **125**(6): p. 653-658.
9. Tveito, A., et al., *Inversion and computational maturation of drug response using human stem cell derived cardiomyocytes in microphysiological systems.* Scientific reports, 2018. **8**(1): p. 1-14.
10. Tveito, A., et al., *Computational translation of drug effects from animal experiments to human ventricular myocytes.* Scientific Reports, 2020. **10**(1): p. 1-11.
11. Sube, R. and E.A. Ertel, *Cardiomyocytes Derived from Human Induced Pluripotent Stem Cells: An In-Vitro Model to Predict Cardiac Effects of Drugs.* Journal of Biomedical Science and Engineering, 2017. **10**(11): p. 527.
12. Navarrete, E.G., et al., *Screening drug-induced arrhythmia using human induced pluripotent stem cell–derived cardiomyocytes and low-impedance microelectrode arrays.* Circulation, 2013. **128**(11_suppl_1): p. S3-S13.
13. Lieu, D.K., et al., *Mechanism-based facilitated maturation of human pluripotent stem cell-derived cardiomyocytes.* Circ Arrhythm Electrophysiol, 2013. **6**(1): p. 191-201.
14. Veerman, C.C., et al., *Immaturity of human stem-cell-derived cardiomyocytes in culture: fatal flaw or soluble problem?* Stem Cells Dev, 2015. **24**(9): p. 1035-52.
15. Tu, C., B.S. Chao, and J.C. Wu, *Strategies for Improving the Maturity of Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes.* Circ Res, 2018. **123**(5): p. 512-514.
16. Blinova, K., et al., *International multisite study of human-induced pluripotent stem cell-derived cardiomyocytes for drug proarrhythmic potential assessment.* Cell reports, 2018. **24**(13): p. 3582-3592.
17. Sala, L., M. Bellin, and C.L. Mummery, *Integrating cardiomyocytes from human pluripotent stem cells in safety pharmacology: has the time come?* British journal of pharmacology, 2017. **174**(21): p. 3749-3765.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18. Alhusseini, M.I., et al., *Machine Learning to Classify Intracardiac Electrical Patterns During Atrial Fibrillation: Machine Learning of Atrial Fibrillation.* Circulation: Arrhythmia and Electrophysiology, 2020. **13**(8): p. e008160.
19. Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural computation, 1997. **9**(8): p. 1735-1780.
20. Ballinger, B., et al. *DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction*. in *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.
21. He, R., et al., *Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM.* IEEE Access, 2019. **7**: p. 102119-102135.
22. Hou, B., et al., *LSTM Based Auto-Encoder Model for ECG Arrhythmias Classification.* IEEE Transactions on Instrumentation and Measurement, 2019.
23. Warrick, P. and M.N. Homsi. *Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks*. in *2017 Computing in Cardiology (CinC)*. 2017. IEEE.
24. Oh, S.L., et al., *Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats.* Computers in biology and medicine, 2018. **102**: p. 278-287.
25. Chen, C., et al., *Automated arrhythmia classification based on a combination network of CNN and LSTM.* Biomedical Signal Processing and Control, 2020. **57**: p. 101819.
26. Wang, L. and X. Zhou, *Detection of congestive heart failure based on LSTM-based deep network via short-term RR intervals.* Sensors, 2019. **19**(7): p. 1502.
27. Bian, M., et al. *An accurate lstm based video heart rate estimation method*. in *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. 2019. Springer.
28. Yildirim, O., et al., *A new approach for arrhythmia classification using deep coded features and LSTM networks.* Computer methods and programs in biomedicine, 2019. **176**: p. 121-133.
29. Wang, E.K., X. Zhang, and L. Pan, *Automatic classification of CAD ECG signals with SDAE and bidirectional long short-term network.* IEEE Access, 2019. **7**: p. 182873-182880.
30. Martis, R.J., et al., *Application of higher order cumulant features for cardiac health diagnosis using ECG signals.* International journal of neural systems, 2013. **23**(04): p. 1350014.
31. Liu, F., et al. *A LSTM and CNN based assemble neural network framework for arrhythmias classification*. in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019. IEEE.
32. Yildirim, Ö., *A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification.* Computers in biology and medicine, 2018. **96**: p. 189-202.
33. Cai, C., et al., *Deep learning-based prediction of drug-induced cardiotoxicity.* Journal of chemical information and modeling, 2019. **59**(3): p. 1073-1084.
34. Zhang, Y., et al., *Prediction of hERG K+ channel blockage using deep neural networks.* Chemical Biology & Drug Design, 2019. **94**(5): p. 1973-1985.
35. Dickson, C.J., C. Velez-Vega, and J.S. Duca, *Revealing molecular determinants of hERG blocker and activator binding.* Journal of chemical information and modeling, 2019. **60**(1): p. 192-203.
36. Ryu, J.Y., et al., *DeepHIT: a deep learning framework for prediction of hERG-induced cardiotoxicity.* Bioinformatics, 2020. **36**(10): p. 3049-3055.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

37. Yang, P.-C., et al., *A computational pipeline to predict cardiotoxicity: From the atom to the rhythm.* Circulation research, 2020. **126**(8): p. 947-964.
38. Li, Z., et al., *General principles for the validation of proarrhythmia risk prediction models: an extension of the CiPA in silico strategy.* Clinical Pharmacology & Therapeutics, 2020. **107**(1): p. 102-111.
39. Tanskanen, A.J. and L.H. Alvarez, *Voltage noise influences action potential duration in cardiac myocytes.* Mathematical biosciences, 2007. **208**(1): p. 125-146.
40. O'Hara, T., et al., *Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation.* PLoS computational biology, 2011. **7**(5): p. e1002061.
41. Kernik, D.C., et al., *A computational model of induced pluripotent stem-cell derived cardiomyocytes incorporating experimental variability from multiple data sources.* The Journal of physiology, 2019.
42. Fawcett, T., *An introduction to ROC analysis.* Pattern recognition letters, 2006. **27**(8): p. 861-874.
43. Powers, D.M., *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.* 2011.
44. Gustafsson, F., *Determining the initial states in forward-backward filtering.* IEEE Transactions on signal processing, 1996. **44**(4): p. 988-992.
45. Fabbri, A., et al., *Required GK1 to suppress automaticity of iPSC-CMs depends strongly on IK1 model structure.* Biophysical Journal, 2019. **117**(12): p. 2303-2315.
46. Atkinson, K.E., *An introduction to numerical analysis*. 2008: John wiley & sons.
47. Li, M., et al., *Overexpression of KCNJ2 in induced pluripotent stem cell-derived cardiomyocytes for the assessment of QT-prolonging drugs.* Journal of Pharmacological Sciences, 2017. **134**(2): p. 75-85.
48. Cheng, J., L. Dong, and M. Lapata, *Long short-term memory-networks for machine reading.* arXiv preprint arXiv:1601.06733, 2016.
49. Olah, C., *Understanding LSTM Networks. Aug. 2015.* URL https://colah. github. io/posts/2015-08-Understanding-LSTMs, 2017.
50. Krogh, A., *What are artificial neural networks?* Nature biotechnology, 2008. **26**(2): p. 195-197.
51. Carugo, O., F. Eisenhaber, and Carugo, *Data mining techniques for the life sciences*. Vol. 609. 2010: Springer.
52. Goodfellow, I., Y. Bengio, and A. Courville, *Deep learning*. 2016: MIT press.
53. Murphy, K.P., *Machine learning: a probabilistic perspective*. 2012: MIT press.
54. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization.* arXiv preprint arXiv:1412.6980, 2014.
55. Zaremba, W., I. Sutskever, and O. Vinyals, *Recurrent neural network regularization.* arXiv preprint arXiv:1409.2329, 2014.
56. Ketkar, N., *Introduction to pytorch*, in *Deep learning with python*. 2017, Springer. p. 195-208.
57. Devore, J.L., *Probability and Statistics for Engineering and the Sciences*. 2011: Cengage learning.
58. Mason, S.J. and N.E. Graham, *Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation.* Quarterly Journal of the Royal Meteorological Society: A journal of the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.28.317461; this version posted September 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

atmospheric sciences, applied meteorology and physical oceanography, 2002. **128**(584): p. 2145-2166.

59. Kernik, D.C., et al., *A computational model of induced pluripotent stem-cell derived cardiomyocytes for high throughput risk stratification of KCNQ1 genetic variants.* PLOS Computational Biology, 2020. **16**(8): p. e1008109.
60. Yang, P.C., et al., *A computational modelling approach combined with cellular electrophysiology data provides insights into the therapeutic benefit of targeting the late Na+ current.* The Journal of physiology, 2015. **593**(6): p. 1429-1442.
61. Yoshinaga, D., et al., *Phenotype-based high-throughput classification of long QT syndrome subtypes using human induced pluripotent stem cells.* Stem cell reports, 2019. **13**(2): p. 394-404.