eLife

Reviewed Preprint
v1 • March 24, 2026
Not revised

Reviewed Preprint
v2 • July 2, 2026
Revised by authors

For correspondence:
ceclancy@ucdavis.edu

Competing interests: DK Lieu serves as a scientific consultant for Novoheart, Ltd.

Funding: See page 20

Reviewing editor: Thomas Hund, The Ohio State University, United States

Large-scale synthetic data enable digital twins of human excitable cells

Pei-Chi Yang[1,2], Mao-Tsuen Jeng[2], Deborah K Lieu[3,4], Regan L Smithers[3,4], Gonzalo Hernandez-Hernandez[1,2], L Fernando Santana[2], Colleen E Clancy[1,2,5]

[1]Center for Precision Medicine and Data Science, University of California, Davis, Davis, United States • [2]Department of Physiology and Membrane Biology, University of California, Davis, Davis, United States • [3]Department of Internal Medicine, Division of Cardiovascular Medicine, University of California, Davis, Davis, United States • [4]Institute for Regenerative Cures, University of California, Davis, Davis, United States • [5]Department of Pharmacology, University of California, Davis, Davis, United States

eLife Assessment

In this **important** study, the authors present an interesting platform for digital twin construction of iPSC-CMs using an AI-based approach. The concept is timely and could have meaningful impact as the field continues to explore integration of computational and experimental models. The evidence is **convincing** overall, although additional attention to framing and calibration of claims would enhance clarity and better reflect the current level of validation.

https://doi.org/10.7554/eLife.110013.2.sa3

Abstract

Individual variability shapes how diseases manifest, how patients respond to therapy, and how rare phenotypes arise. Conventional experimental approaches obscure variation by averaging, which limits mechanistic insight and predictive accuracy. We present a computational framework that builds digital twins of human-induced pluripotent stem cell-derived cardiomyocytes from a single optimized voltage clamp experiment. The framework depends on massive synthetic datasets comprising simulated cells that span broad ionic and electrophysiological ranges. These synthetic data make it possible to control parameters precisely, explore biological variability comprehensively, and train models beyond the limits of experimental data. A neural network trained on synthetic data then inferred biophysical parameters from experimental recordings from live cells, reproducing distinct electrophysiological features. Our study unites computational modeling, data simulation, and learning to enable scalable, precise, individualized cardiac electrophysiology modeling and can be readily extended to any electrically active cell type.

Introduction

Individual variability determines the emergence and severity of disease, as well as individual therapeutic response in the context of both inherited and acquired conditions[1–3]. Traditional preclinical models are designed to minimize variability to isolate specific biological effects. While this reductionist approach simplifies experimental interpretation, it fails to reflect the heterogeneity that exists in human populations. The consequence is a persistent translational gap, an incomplete mechanistic understanding of disease, and a higher risk of adverse drug reactions[4–8]. Digital twins constitute a mechanistic approach to integrate individualized data with simulation-based inference to predict physiology and therapeutic outcomes at the level of an individual cell[9–12].

Human induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) are a powerful platform for studying electrophysiology in a patient-specific context[13–19]. However, experimental measurements of iPSC-CM action potentials (APs) or ionic currents provide only partial information about the underlying mechanisms[20–24]. The relationship between measured signals and the large number of biophysical parameters governing ion channel kinetics, calcium handling, and membrane transport is complex and nonlinear, making direct parameter estimation from small experimental datasets impossible. Existing fitting approaches are slow, require extensive specialized expertise, and are not scalable to capture individual variability[25–31].

Mathematical models of cardiac electrophysiology comprise a complementary framework to link cellular dynamics and underlying ionic mechanisms[32,33]. Personalized digital twins have not yet been realized in part because of ongoing challenges to parameterize models accurately and quickly from cell-specific experimental data. Recent advances in deep learning now make it possible to solve the associated complex inverse problem by deriving the mappings from high-dimensional time-dependent recordings to the underlying biophysical parameters. The key is the use of massive synthetic datasets from large *in silico* cell populations to generate the training data necessary for accurate, generalizable inference. It is not possible to develop a comparable process with purely experimental data.

Here, we present an integrated computational modeling and simulation, deep learning framework, and experimental recordings demonstrating the generation of high-fidelity, electrophysiology-based digital twins of iPSC-CMs from a single optimized voltage-clamp protocol. Large-scale synthetic model iPSC-CM populations were used to train a fully connected neural network to predict 52 biophysical parameters governing six major ionic currents. A deep learning-guided optimization loop was used to design voltage-clamp protocols that maximize information content for parameter recovery. The primary goal of this framework is to enable population-level interpretation of electrophysiological variability by mapping experimental recordings to underlying ionic parameters. The models reconstructed from the inferred parameters reproduce the AP waveform, calcium dynamics, and the contributions of individual ionic currents. When applied to experimental data, this approach generates high-fidelity digital twins of experimental recordings with high accuracy across temperatures and morphological variability.

This work addresses a major translational challenge by providing a scalable, temperature-agnostic, morphology-inclusive method for personalized cardiac modeling to bridge the gap between experimental recordings and mechanistic insight. We present a framework for predictive digital twin applications in precision cardiovascular medicine and beyond, since the method can be applied to any electrically active cell type. By uniting protocol optimization, large-scale synthetic training, and deep learning-based inference, this work advances the field toward scalable digital twin technology with broad potential for basic research, preclinical testing, and personalized cardiovascular medicine.

Results

Variability and Digital Twin Modeling of iPSC-CMs

To assess the electrophysiological variability inherent in populations of iPSC-CMs, we first generated a simulated population of 200,000 spontaneously beating cells by applying ±40% random variation to 52 biophysical model parameters (see Supplementary Information). These parameters govern six key ionic currents central to cardiac AP generation and repolarization: I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f. Examples of population variability are shown in 20 APs in Figure 1A. With a broad range of parameter perturbations, extensive variability was observed in the model population, as seen in experimental iPSC-CM preparations.

The simulations revealed a wide distribution of AP morphologies and spontaneous firing frequencies (examples in Figure 1B). APs varied substantially in upstroke velocity, plateau duration, and repolarization slope, with some cells exhibiting prolonged plateaus while others

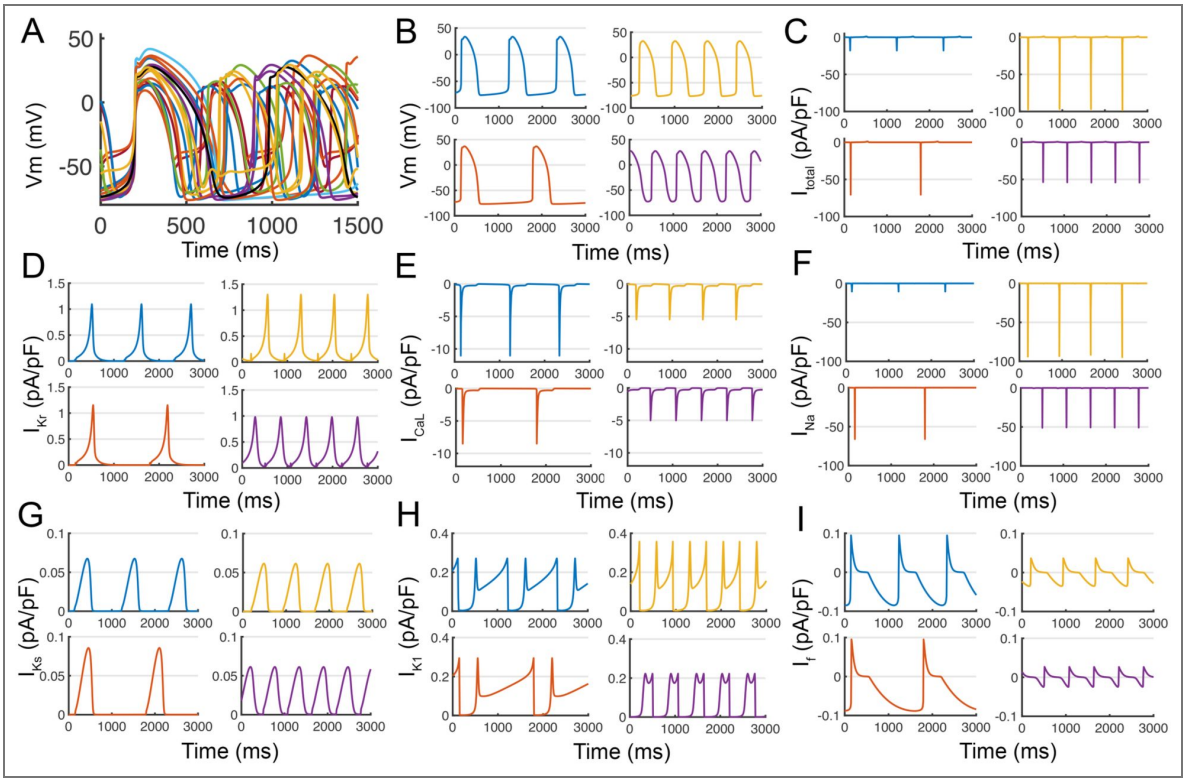


Figure 1. Huge variability captured in simulated iPSC-derived cardiomyocyte populations.

(A) To illustrate population variability, 20 action potentials (APs) were shown, each resulting from ±40% random variation applied to 52 parameters governing six key ionic currents (I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f) in the baseline Kernik iPSC-CM model [44], within a simulated population of 200,000 spontaneously beating cells. **(B)** These perturbations yielded a wide spectrum of APs, with substantial variation in both waveform and frequency. The corresponding total ionic current **(C)** and its decomposition into individual current components (**D–I**) are shown.

demonstrated rapid repolarization and shorter cycle lengths. Differences in spontaneous beat rate reflected variability in diastolic depolarization rates, which were closely linked to variations in I_f amplitude and I_{K1} strength across the simulated population.

To dissect the mechanistic drivers of electrophysiological diversity, we examined the total membrane current (I_{Total}) shown in Figure 1C and decomposed it into individual ionic components for representative cells across the population (Figure 1D–I). The heterogeneity in total current waveforms was accompanied by diverse current profiles across I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and $I_{f.}$ Some cells exhibited larger I_{CaL} amplitudes and prolonged calcium entry during the plateau, while others showed enhanced I_{K1-} or I_{Kr}-mediated repolarizing currents. Differences in I_{Na} upstroke magnitude and I_f driven diastolic depolarization contributed to variation in excitability and cycle length. Together, these simulations underscore the high degree of functional variability that can emerge from relatively modest variation in channel kinetics and conductance properties. The synthetic cell population mirrors the experimental heterogeneity observed in iPSC-CM populations and underscores the likely importance of accounting for cellular variability in predictive modeling and ultimately, in individual drug response evaluation.

We next developed a deep learning-based workflow to generate fully parameterized models of iPSC-CM electrophysiology from whole-cell current data obtained with one simple voltage-clamp protocol (Figure 2). To create the training dataset, we generated a large synthetic iPSC-CMs population by introducing random variations to 52 biophysical parameters governing six key ionic currents in the baseline model: I_{Kr}, I_{CaL}, I_{Na}, $I_{Ks,}$ $I_{K1,}$ and I_{f}. The parameter perturbations produced an array of simulated APs and whole-cell currents, reflecting broad physiological variability (**left**).

As part of the workflow, a voltage-clamp protocol (black trace in left middle) was applied to each synthetic cell to generate complex current responses to dynamic membrane potential input. The protocol was designed to sequentially engage major ionic currents in the synthetic cell, ensuring that the training data contained informative current signatures for each underlying conductance. The resulting total current traces (I_{Total}) served as inputs to a fully connected neural network (red trace in **left middle)**.

The network was trained to map the raw current inputs to the 52 underlying model parameters, with training performance evaluated by mean squared error (MSE) on held-out test datasets. An example parameterization for the fast sodium current (I_{Na}) illustrates that the network infers gating kinetic parameters and maximal conductance (**right**).

Deep Learning Approaches for Voltage Clamp Optimization and Parameter Estimation

To maximize the information content of experimental recordings for parameter inference, we developed a deep learning-guided strategy to design a computationally optimized voltage-clamp protocol (Figure 3). Starting from a broad set of candidate voltage steps, the algorithm iteratively evaluated the MSE between predicted and true parameters across large synthetic iPSC-CM populations. In each cycle, the most informative test potential was selected and incorporated into the evolving protocol, while less informative steps were discarded.

Each cycle began with a -100 mV holding potential for 250 ms, followed by systematic sweeps of test potentials from -120 mV to +50 mV in 10 mV increments. A population of 200,000 computational iPSC-CMs with randomized parameter sets was simulated under each protocol iteration, and the resulting whole-cell currents were used to train the deep learning model to predict all 52 ionic parameters. The potential producing the lowest MSE was then selected as the optimal step and applied for 250 ms before the sequence resumed. This iterative optimization loop was repeated every 7000 ms, with a transient step to −120 mV at 6000 ms to probe hyperpolarization-activated channels.

This adaptive process converged on a composite voltage-clamp waveform that effectively exposed the amplitude and kinetics of all six major ionic currents while substantially reducing voltage variability through the protocol duration. The optimized sequence preserved temporally distinct current signatures for I_{Kr}, I_{CaL}, I_{Na}, $I_{Ks,}$ $I_{K1,}$ and $I_{f,}$ enabling robust parameter estimation.

Figure 2. **Digital twins of human iPSC-CMs from one simple voltage-clamp recording.**

A large population of synthetic iPSC-CMs (left) is generated by introducing variation to 52 biophysical parameters governing key ionic currents in the baseline model: I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f (*highlighted with red asterisks* in the schematic on the right). A computationally optimized voltage-clamp protocol (black trace, left middle) is applied to generate a distinct whole-cell current (I_{Total}, red trace), enabling cell-wide excitation of key ion channels. The simulated whole-cell currents I_{Total} from large synthetic datasets serve as inputs to a fully connected neural network trained to map raw current responses to the 52 underlying model parameters. The deep learning model is trained to predict optimized parameter values by inferring gating kinetics and maximal conductance for each ionic species. An example formulation for the fast sodium current (I_{Na}) is shown, with inferred parameters (x_1–x_5) contributing to gating and conductance (right).

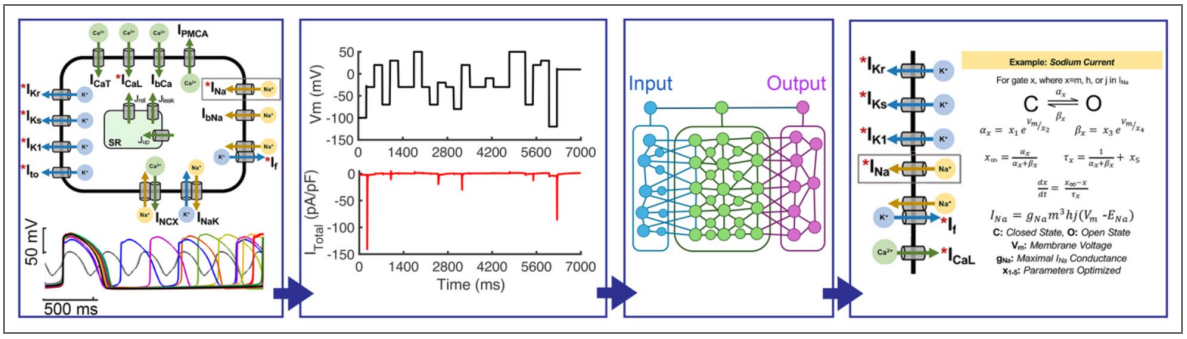
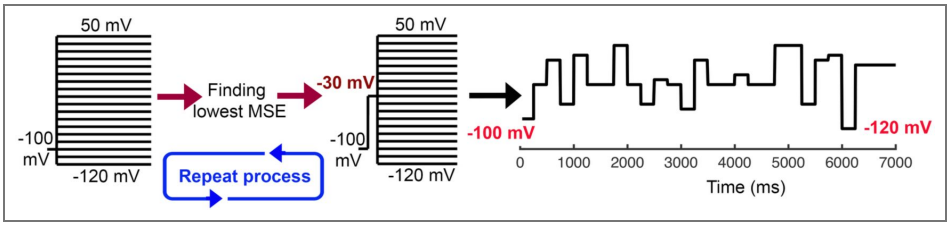


Figure 3. **Deep learning guided optimization of a voltage clamp protocol.**

Each iteration began with a −100 mV holding potential for 250 ms, followed by sequential testing potentials from −120 mV to +50 mV in 10 mV increments. A total of 200,000 synthetic samples were generated per training cycle. The optimal testing potential, identified by the lowest MSE from the deep learning model, was then applied for 250 ms. This optimization cycle repeated every 7000 ms. At 6000 ms, the potential was transiently stepped to −120 mV for 250 ms before resuming the next testing potential sequence. The schematic illustrates the iterative loop of model evaluation, MSE-based selection, and protocol updating, leading to the optimized composite voltage.



When applied to held-out synthetic iPSC-CMs test data, the optimized voltage clamp protocol (Figure 4A, left) produced digital whole-cell currents (Figure 4A, right) that were effective as network training data, confirming that the protocol provided sufficient information for accurate model personalization. Training data size was still influential and determined predictive accuracy, with improvements in trained networks across datasets containing 1,000, 10,000, and 200,000 synthetic cells (Figure 4B). For the smallest dataset (1,000 cells), the test error plateaued at an MSE of 0.0036, with a median mean absolute error (MAE) of 0.041 across parameters. The test error curve exhibited an asymmetric U-shape across epochs, suggesting overfitting to the limited training set (**4B, left**). Increasing the dataset size to 10,000 cells improved test performance, reducing the median MAE to 0.038 (a 7.3% decrease) and narrowing the error distribution (**4B, middle**). The largest dataset of 200,000 cells had very low prediction errors, with training and test curves converging to a median mean absolute error (MAE) of 0.02 (**4B, right**). Again, the optimized protocol resulted in corresponding MAE distributions that narrowed substantially with increasing dataset size, indicating improved stability and generalizability of parameter predictions.

The deep learning network-generated parameters led to improved model waveforms (red) that duplicated ground truth simulated data (blue) for all outputs, with high fidelity in both amplitude and temporal features (Figure 4C). We observed overlap of AP morphology, Ca^{2+} transient dynamics, and the activation and inactivation profiles of each ion channel. These findings demonstrate that large, diverse training datasets are essential for high-fidelity parameter inference, enabling digital twins to accurately replicate both the quantitative dynamics and qualitative features of iPSC-CM electrophysiology. This scalability positions the framework for reliable application to experimental datasets where precise recovery of ionic properties is critical.

Digital twins of human iPSC-CMs reproduce experimental electrophysiology and reveal intrinsic variability in drug response

While we have shown that the deep learning-based approach can create high-fidelity digital twins when trained and tested with in silico models of iPSC-CMs, the ultimate proving ground for the digital twin framework is application to real human iPSC-CM recordings. Therefore, we applied the optimized voltage-clamp protocol in its most demanding and key test to derive digital twins from real cells (Figure 5).

The experimental cells in this study exhibited remarkably large variability in AP morphology, cycle length, and plateau characteristics, reflecting the heterogeneity typical of iPSC-CM preparations.

Capturing this breadth of behavior in the modeling pipeline required a correspondingly broad synthetic training dataset. To achieve this, we generated a large *in silico* population of 1,100,000 iPSC-CMs by applying ±40% random variation to 52 biophysical parameters governing six major ionic currents I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f relative to the baseline model. This expanded range ensured that extreme and rare phenotypes present in the experimental recordings were represented in the training space.

A second major consideration was temperature. While the iPSC-CM models were originally parameterized at 37 °C, the experimental data from iPSC-CM are commonly acquired at room temperature. To reconcile this mismatch, we systematically slowed all channel kinetics in the synthetic training population using a Q_{10} factor of 3, a coarse but effective first approximation, to estimate the slower gating dynamics observed experimentally. This adjustment allowed the network to be trained on current waveforms that more closely matched the temporal characteristics of the experimental recordings.

Whole-cell currents recorded from two live iPSC-CMs (Figure 5A and B, top) in response to the protocol in Figure 3 served as the sole input to the trained deep learning network, which successfully inferred all 52 biophysical parameters associated with six major ionic currents, I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f. These parameter sets were used to instantiate personalized digital twins for each experimental cell. Simulated APs (**A, B, bottom**) from the synthetic population at room

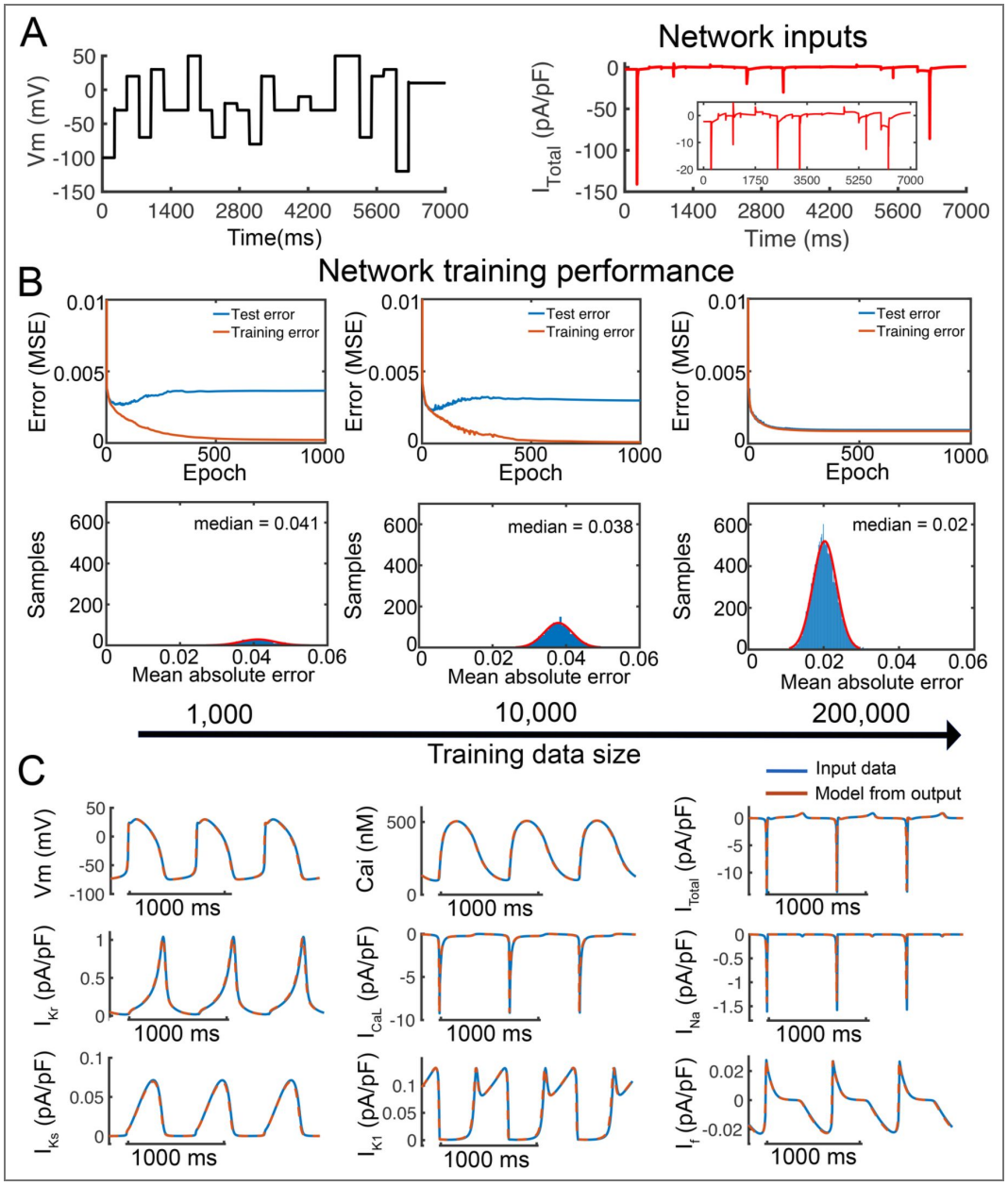


Figure 4. Massive synthetic data to train and test a deep learning algorithm for ion channel parameter estimation.

(A) A deep learning-derived optimized voltage-clamp protocol (left) was designed to activate a broad set of ionic currents in iPSC-CMs by applying dynamic membrane potential steps (black trace), producing distinctive whole-cell current responses (I_{Total}, red trace). The inset highlights fine-scale current kinetics captured during the protocol. These I_{Total} serve as network inputs for parameter inference. **(B)** Prediction accuracy for individual model parameters was evaluated using training datasets of 1,000, 10,000, and 200,000 synthetic cells (left to right), generated with a ±20% parameter perturbation range. For the smallest dataset, the test error exhibited an asymmetric U-shaped curve across training epochs, indicating limited generalizability. As the dataset size increased, training and test errors converged, with median mean absolute errors (MAE) of 0.041, 0.038, and 0.020 for the three dataset sizes, respectively. Bottom panels show MAE distributions across parameters, demonstrating progressively narrower error ranges as training datasets grow. **(C)** APs (V_m), intracellular calcium transients (Cai), total ionic current (I_{Total}), and six major individual ionic currents (I_{Kr}, I_{CaL}, I_{Ks}, I_{K1}, I_{Na}, I_f) were simulated from the predicted parameters of a single test cell using the deep learning network trained on 200,000 samples. Blue traces represent the original simulated data (input), and red traces show the model outputs. The close overlap confirms faithful reproduction of cell-specific electrophysiological behavior. See Figure S1 for a detailed comparison of the corresponding parameter values.

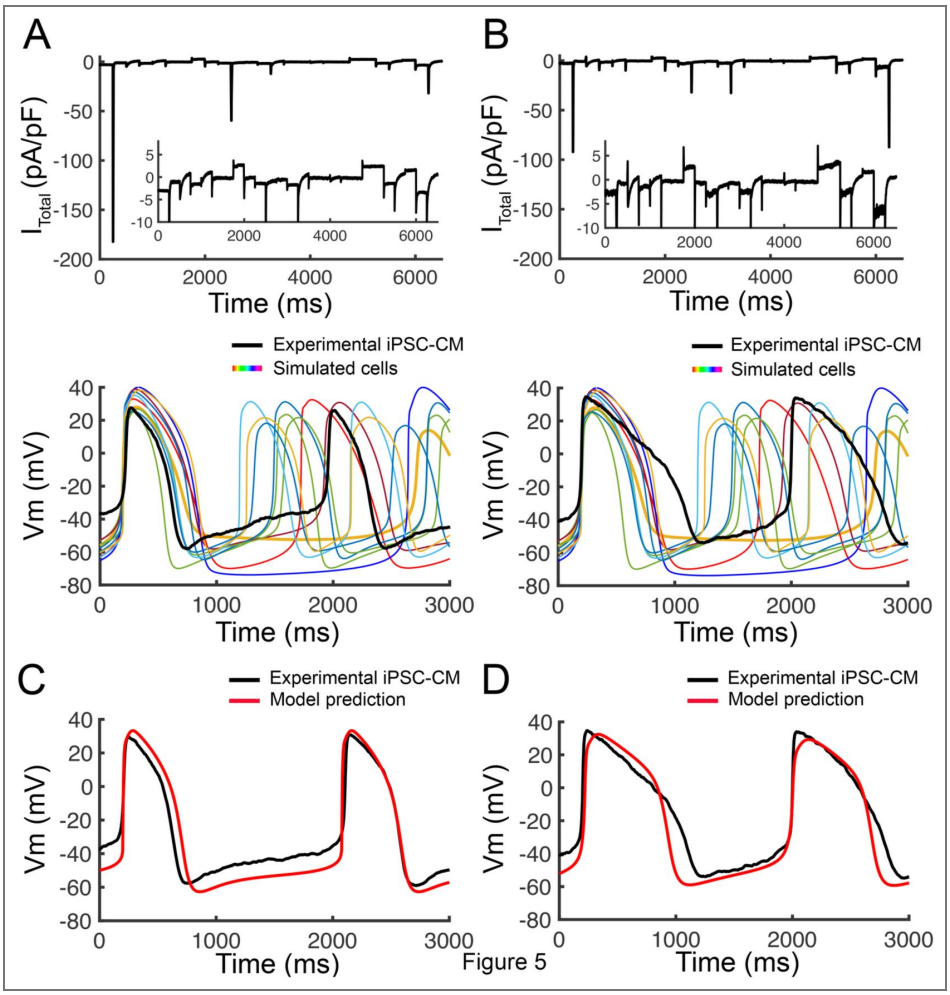


Figure 5. Digital twin generation and predictive modeling from real cells in iPSC-CM experimental recordings.

(A, B, top) Whole-cell currents from two live iPSC-CMs in response to the optimized voltage-clamp protocol (Figure 3) were used as input to the trained deep learning network. Insets show expanded views of the responses. **(A, B, bottom)** Ten simulated APs from a synthetic population of 1,100,000 iPSC-CMs at room temperature were generated by introducing ±40% random variation to 52 biophysical model parameters from baseline, governing six major ionic currents: (I_{Kr}, I_{CaL}, I_{Ks}, I_{K1}, I_{Na}, I_f). Colored traces represent exemplar simulated cells from training data; overlaid black traces are experimentally recorded APs from two representative iPSC-CMs of the human iPSC cell line iPS-6-9-9T.B. (**C, D**) Digital twin models of the same experimental iPSC-CMs shown in panels A and B. Digital twins were created by extracting all 52 model parameters from the experimental whole-cell current using the deep learning inference pipeline. These parameters were used to instantiate cell-specific computational models, and AP simulations (red traces) were generated. The close overlay between the experimental traces (**black**) and digital twin predictions (**red**) demonstrates the success of the framework to accurately derive cell-specific digital twins from real cells.

temperature reflect the range of electrophysiological variability observed in experimental iPSC-CMs, indicating that the training data adequately capture the diversity of the experimental recordings.

When simulated, the digital twins reproduced the experimentally recorded APs with excellent agreement, capturing fine-scale features of depolarization rate, plateau amplitude and duration, and repolarization slope (Figure 5C–D), confirming that the pipeline reconstructs the full mechanistic basis of electrophysiology from a single experimental protocol.

To investigate how intrinsic variability among iPSC-CMs influences proarrhythmic susceptibility, we compared the responses of two representative model cells with distinct repolarization dynamics (Figure 6A–B). Under control conditions at 37 °C, both **Cell 1** and **Cell 2** generated stable APs with similar morphology. Exposure to the selective rapid delayed rectifier potassium current I_{Kr} blocker E-4031 (50 nM) markedly prolonged repolarization in both models, but only **Cell 1** exhibited early afterdepolarizations (EADs), whereas Cell 2 maintained stable repolarization. These findings highlight the strong impact of intrinsic cellular variability on proarrhythmic susceptibility and emphasize the value of digital twin models in capturing the heterogeneity of proarrhythmic responses to drug exposure across individual cells.

To extend this analysis beyond single cells, we generated two *in silico* populations (n = 4000) derived from the network-predicted parameter sets of **Cell 1** and **Cell 2** (Figure 6C). Each population incorporated ±20% perturbations to all parameters governing six major ionic currents (I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f), capturing the intrinsic cell-to-cell variability expected within a single human iPSC line. Under control conditions, both populations showed stable spontaneous firing with overlapping distributions of AP, reflecting baseline heterogeneity in electrophysiological properties.

The resulting population captured a broad distribution of repolarization phenotypes (mean APD_{90} = 403 ± 47 ms). Upon exposure to increasing concentrations of E-4031, the incidence of EADs increased in a dose-dependent manner, reaching approximately 3% of simulated cells at 50 nM (Figure 6D), based on pooled variants from both **Cell 1** and **Cell 2**. Although this represents a relatively small fraction of the population, it reflects the emergence of a susceptible subpopulation arising from intrinsic electrophysiological variability. These results demonstrate that modest cell-to-cell variability in ion-channel properties can substantially amplify heterogeneity in proarrhythmic responses within a population. The digital-twin framework, therefore, enables statistically robust quantification of arrhythmia risk by bridging single-cell electrophysiological diversity to population-level drug sensitivity.

Discussion

This study demonstrates, for the first time, the development of digital twins of human iPSC-CMs directly from a single experimental recording that is readily generalizable to any excitable cell type. By combining an optimized voltage-clamp protocol, large-scale synthetic training populations, and deep learning–based parameter inference, we reconstructed the complete ionic parameter set underlying individual cell electrophysiology and reproduced AP and current waveforms with high fidelity. This study addresses a longstanding limitation in cardiac modeling: the inability to rapidly and accurately resolve the individualized ionic mechanisms that drive variability in both health and disease. The framework proved robust to two of the most challenging sources of variation in iPSC-CM studies, temperature differences and broad waveform diversity, highlighting the potential for a scalable and adaptable tool for basic research, preclinical safety testing, and precision medicine.

A broad range of AP morphologies and firing rates in our simulated iPSC-CM population were readily evoked with changes of ±40% across ionic parameters (Figure 1). This mirrors experimental reports of substantial heterogeneity in iPSC-CM electrophysiological properties, which may arise from differences in differentiation state, ion channel expression, and culture

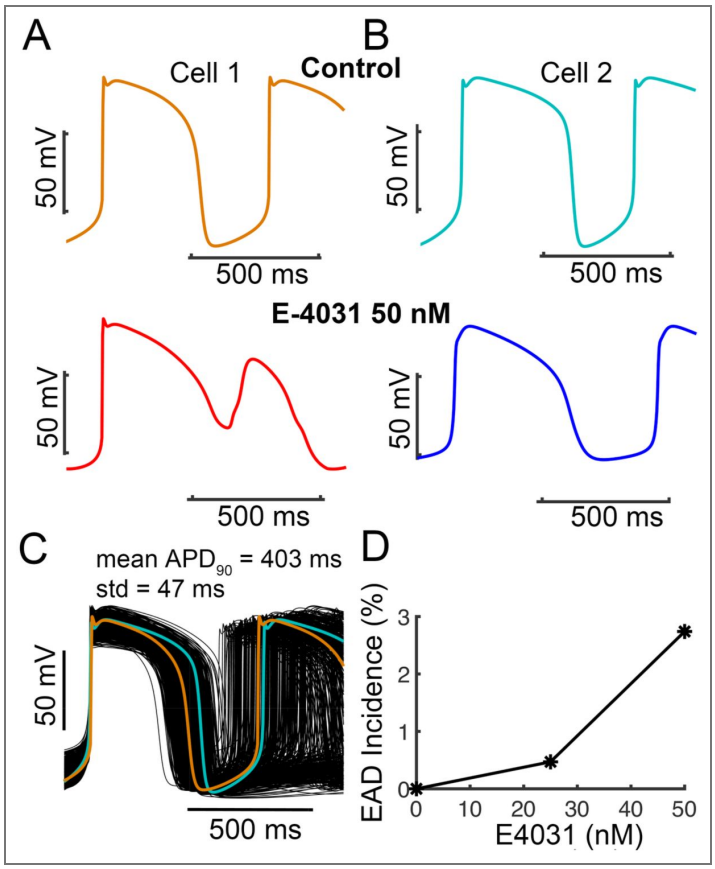


Figure 6. Cell-to-cell variability drives population-level responses to drug application in large synthetic iPSC-CM digital twins.

(A) iPSC-CM model **Cell 1** (corresponding to Figure 5A). **(B)** iPSC-CM model **Cell 2** (corresponding to Figure 5B). In both panels, the **top traces** show control simulations at physiological temperature (37 °C). The **middle row** shows responses in the presence of E-4031 (50 nM). EADs occurred in **Cell 1** (red) at 50 nM, but not in **Cell 2** (blue). **(C)** We applied ±20% perturbations to all parameters governing six key ionic currents (I_{K1}, I_{Kr}, I_{Ks}, I_{CaL}, I_{Na}, I_f) in Cell 1 (orange) and Cell 2 (cyan) to generate a population of virtual cells (n = 4000), capturing the full spectrum of cell variability within a cell line. Overlaid membrane potential traces (black) show APs from Cell 1 and Cell 2. The population average action potential duration at 90% repolarization (APD_{90}) was 403 ± 47 ms. **(D)** Incidence of EADs (%) in the virtual population as a function of E-4031 concentration.

conditions[34–39]. Such diversity has important implications for both mechanistic interpretation and predictive modeling. It suggests that population-based approaches, rather than reliance on single-cell models, are essential for capturing the spectrum of cellular responses[40–43].

The electrophysiological phenotype of the human iPSC-CMs in this study is represented using the Kernik–Clancy iPSC-CM model [44], which was developed to capture the range of AP morphologies and spontaneous behaviors observed experimentally within a single cell line [37]. This population-based modeling framework incorporates variability in key ionic conductances to reproduce the spectrum of depolarization, repolarization, and automaticity characteristics commonly reported in human iPSC-CMs, while constraining parameter ranges to avoid non-physiological phenotypes.

Within this framework, we focused on six primary ionic currents that are well established as dominant determinants of AP shape and variability in iPSC-CMs. These currents govern critical features, including upstroke velocity, plateau dynamics, repolarization, and spontaneous activity, and their modulation has been shown to recapitulate experimentally observed cell-to-cell heterogeneity [23,38,39,45–52]. Although many ionic currents are developmentally regulated and may vary across human iPSC-CM lines [53], this reduced set provides a parsimonious representation that captures the expected electrophysiological variability within a given cell line.

Importantly, the model is designed to be extensible. Additional currents or alternative parameterizations can be incorporated to reflect differences in maturation state, donor background, or differentiation protocol, enabling future studies to explicitly address variability across human iPSC-CM sources while maintaining consistency with the underlying framework.

Figure 2 highlights the key technology developed in this study, a deep learning-based workflow to recover the key ionic parameter set of a cell from a single voltage-clamp experiment. The synthetic training population of iPSC-CMs, generated by introducing random variation to 52 parameters governing six key ionic currents I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f, ensured that the network was exposed to a broad spectrum of AP morphologies and whole-cell current profiles.

We developed a data-driven protocol design process guided by deep learning, as shown in Figure 3. In this approach, candidate voltage segments were iteratively evaluated to minimize parameter prediction error in synthetic populations, enabling the discovery of less variable, more targeted protocols. This optimization strategy effectively probed multiple ionic currents while reducing voltage variation, thereby improving voltage control and cell tolerance. The adaptive nature of the design process ensured that protocols were both information-rich and experimentally viable, providing a robust foundation for accurate parameter inference and digital twin construction from live iPSC-CM recordings. The iterative deep learning-guided approach to voltage-clamp protocol optimization offers clear advantages over traditional manual design strategies. Conventional protocols are typically based on heuristic knowledge of channel gating properties and are rarely tailored to maximize information for multi-parameter inference. Our adaptive framework directly couples experimental design with model performance metrics, using MSE as a quantitative feedback signal to iteratively refine the protocol. This ensures that each iteration selectively enriches the data with features most discriminative for parameter recovery, thereby accelerating convergence to an optimal waveform. Importantly, because the optimization is performed *in silico* using large-scale synthetic populations, the resulting protocol is both data-driven and generalizable, capturing a wide range of ionic phenotypes. Once developed, such a protocol can be implemented in experimental systems without modification, offering a powerful route to improve the precision and efficiency of digital twin construction in human cardiomyocytes.

The clear improvement in predictive accuracy and stability with increasing training dataset size, as shown in Figure 4, highlights a fundamental consideration for applying deep learning to biophysically detailed cardiac modeling. Larger datasets capture a wider range of morphological and kinetic variability in ionic currents, reducing the risk of model overfitting and enabling more reliable generalization to new experimental data. The largest dataset produced digital twins that

match ground truth across APs, calcium dynamics, and individual currents, and indicates the value of comprehensive, diverse training data to ensure high-fidelity reproduction at the single-cell level.

The deep learning-guided protocol optimization framework described here lays the groundwork for the capstone result presented in Figure 5, where fully parameterized digital twins reproduced experimental iPSC-CM APs with excellent fidelity. By training on large-scale, information-rich synthetic datasets generated through optimized protocols, the platform achieves robustness to both temperature variation and broad morphological diversity in electrophysiological signals. This adaptability ensures that the same modeling pipeline can be rapidly retuned to diverse experimental conditions and phenotypes without sacrificing accuracy. As a result, the approach is not limited to *in vitro* modeling but is positioned for translational deployment, where personalized, physiology-grounded digital twins could guide patient-specific diagnostics, risk assessment, and therapy optimization. To move seamlessly from protocol design to clinically relevant predictive modeling represents a decisive step toward the integration of digital twins into the practice of precision cardiovascular medicine.

Figure 6 illustrates how intrinsic electrophysiological variability among iPSC-CMs can profoundly influence responses to pharmacological perturbation, even among cells derived from the same genetic background. Small differences in ionic conductance and gating kinetics were sufficient to determine whether a cell remained stable or developed EADs under identical drug exposure.

At the population level, the dose-dependent emergence of EADs provides a quantitative biomarker of proarrhythmic risk that cannot be captured by single deterministic simulations. Importantly, this emergent variability does not require extreme outlier phenotypes but arises naturally from multivariate perturbations around experimentally constrained parameter sets. This highlights the predictive value of large-scale digital twin populations for mechanistic safety pharmacology, enabling the translation of cell-level uncertainty into statistically robust risk metrics. Beyond E-4031, this framework can be generalized to compare patient-derived iPSC-CM lines or therapeutic perturbations, supporting the development of precision cardiotoxicity assessment pipelines that account for both genetic and stochastic sources of electrophysiological diversity.

The success of our parameter inference framework relies critically on the use of large, diverse synthetic datasets that could not be obtained through experimental recordings alone. Training the deep learning model required hundreds of thousands to millions of cellular signals. This is not possible to do experimentally as it would be prohibitively time-consuming, costly, and technically unfeasible to acquire high-fidelity voltage-clamp data at scale. In contrast, computationally generated data allow complete control over underlying parameters, precise labeling of ground truth values, and systematic exploration of parameter space well beyond what is directly observable in experiments. This approach not only accelerates model training but also ensures exposure to rare or extreme phenotypes and improves network generalization across diverse experimental cells.

In terms of scope, there are some limitations to our approach. The framework emphasizes six major ionic currents and a set of supporting mechanisms; however, it does not yet incorporate all ionic currents relevant to cardiac electrophysiology. Expanding the framework to include additional currents would necessitate introducing more parameters, which would increase both model complexity and the scale of training data required, although future iterations will undoubtedly address this limitation. Another limitation is the substantial computational resources needed to generate large synthetic datasets and train deep neural networks. Once training is complete, however, parameter inference is rapid and efficient, making the framework practical for downstream applications.

While developing and validating the digital twin framework *in silico* and *in vitro* using human iPSC-derived cardiomyocytes, the deep learning–guided protocol optimization framework can be extended to a wide range of excitable cell systems and experimental contexts. In principle, the same adaptive feedback loop could be deployed in patch-clamp or multi-electrode array

recordings from any animal or human cell types, including neurons, smooth muscle cells, or engineered tissue constructs, where capturing rich ionic dynamics is essential for mechanistic modeling. Furthermore, by incorporating additional physiological variables such as calcium handling or mechano-electric feedback, the approach could evolve into a multi-modal protocol design tool capable of constraining both electrical and mechanical parameters in digital twin models.

This study examines a single iPSC-CM line across two independent differentiation batches, capturing cell-to-cell variability within each batch. While variability across lines, donors, and protocols was not directly evaluated here, the deep learning framework is structured to accommodate these additional sources of variation in future applications. The synthetic population was designed to span a broad range of physiologically plausible behaviors, providing a useful platform for probing potential response variability. While in this study we did not intend to represent a measured biological distribution, the framework can be readily extended for this purpose. The framework enables scalable prediction of population-level drug responses, and ongoing experimental efforts with larger and more diverse cell samples will further strengthen and refine the model predictions. Importantly, the approach is inherently extensible, and future studies incorporating replicate recordings across multiple cells, lines, and conditions will help deepen understanding of variability and expand its applicability.

We anticipate that near-future studies using the digital twin framework will enable high-throughput single-cell electrophysiological phenotyping for disease modeling, allowing investigators to map patient-derived cardiomyocyte function to individualized ionic mechanisms using a single experiment. The framework maintains accuracy across changes in temperature and accommodates wide variation in AP morphology, making it suitable for multicenter studies in which experimental conditions differ to improve reproducibility and enable harmonization of data. In preclinical safety assessment, temperature-independent digital twins could be used to evaluate the effects of new compounds at scale across a spectrum of cardiac phenotypes, identifying subgroups at increased risk. Looking further ahead, integration of this modeling platform with longitudinal patient data, including genetic information, imaging, and clinical outcomes, could give rise to living cardiac digital twins that evolve over time. Such models may ultimately support predictive diagnostics, optimization of therapy, and continuous monitoring in personalized healthcare.

Methods

Human iPSC-derived Ventricular Cardiomyocyte Differentiation and Culture

The human iPSC line iPS-6-9-9T.B was purchased from WiCell and cultured in StemMACS iPS-Brew XF medium (Miltenyi Biotec) on hESC-qualified Matrigel (Corning) at 37°C with 5% CO_2 and passaged at ~70% confluency. Two independent batches of ventricular cardiomyocytes (VCMs) were differentiated in a 12-well plate by treating confluent human iPSCs with 6 μM CHIR99021 (Tocris) from day 0 to 1 (D0–D1) and 5 μM IWR-1 (Tocris) from day 2 to 4 (D2–D4) in RPMI 1640 supplemented with B27 without insulin (Thermo Fisher). From day 7 onward, human iPSC-derived VCMs were maintained in cardiomyocyte (CM) maintenance medium consisting of RPMI 1640 with B27 supplement with insulin. Cells were pooled from three wells of a 12-well plate and replated at approximately day 14 (D14) and maintained in CM maintenance medium supplemented with 2 μM CHIR99021 for an additional ~7 days. Subsequently, cells were replated onto Matrigel-coated glass coverslips in CM maintenance medium for patch-clamp experiments. Patch-clamp recordings were performed on approximately day 21-25 (D21-25) post-differentiation. Recordings were obtained with cells sampled across 6 coverslips per batch under consistent experimental conditions.

Experimental electrophysiology recordings

All electrophysiological recordings were performed at room temperature with an Axopatch 200A amplifier coupled to an Axon Digidata 1550B plus HumSilencer digitizer (Molecular Devices, San José, CA, USA). Cultured ventricular myocytes were placed on the stage of an inverted microscope (IX-71; Olympus Corporation, Tokyo, Japan) and continuously perfused with Tyrode's solution. Patch pipettes were pulled from borosilicate glass capillaries (Sutter Instrument, Novato, CA, USA) and filled with the appropriate internal solution. The pipette solution contained (in mmol L^{-1}): 140 KCl, 5 NaCl, 10 EGTA, 5 Mg-ATP, and 10 HEPES, adjusted to pH 7.2 with KOH. Ionic currents were recorded in the whole-cell configuration of the patch-clamp technique at a sampling frequency of 10 kHz. After achieving the whole-cell configuration, the amplifier was switched to current-clamp mode for membrane potential measurements. The computationally optimized voltage-pulse protocols (see below) were used to efficiently probe membrane currents. For experimental validation datasets, any incomplete or failed recordings were excluded prior to analysis. Recordings were excluded if they exhibited no spontaneous firing, abnormally slow firing rates, or failed to capture a complete action potential waveform. These criteria were applied consistently across all recordings.

Computationally optimized voltage-clamp protocol

A computationally optimized voltage-clamp protocol was designed to activate individual ionic currents while capturing their dynamic interactions (see Figure 3) sequentially and selectively. This protocol was applied to each synthetic cell to produce time-series data of total ionic current (I_{Total}). Data were generated at a 0.1 ms time step and down-sampled to match the experimental acquisition rate. Following the derivation of the final optimized protocol, the time-series data of I_{Total} current was employed as input to the fully connected neural network, and the output layer contained 52 units corresponding to the model parameters, enabling the extraction of the model parameters.

iPSC-CM electrophysiological model

A baseline iPSC-CM electrophysiological model was parameterized with 52 biophysical variables (see model equations in Supplementary Information) representing maximal conductances, gating kinetics, and calcium handling parameters for six major ionic currents (I_{K1}, I_{Kr}, I_{Ks}, I_{CaL}, I_{Na}, I_f). We selected ionic currents based on the Kernik-Clancy iPSC-CM model from our earlier research [44], which aims to simulate electrophysiological variability observed experimentally using a population-based approach. In this model, variability in key ionic currents was sufficient to generate the diverse action potential morphologies and spontaneous activities seen in iPSC-CMs, while avoiding non-physiological phenotypes. Using this approach, we concentrated on six primary ionic currents that critically influence action potential behavior and variability.

Fully connected layers

Fully connected deep neural networks were implemented in TensorFlow using 26 hidden layers (1024 nodes per layer) with hyperbolic tangent (tanh) activation functions. Each layer performs a linear transformation followed by a nonlinear activation [54,55]:

$$\mathrm{a}^{(k)} = f\left(Z^{(k)}\right) \tag{1}$$

$$Z^{(k)} = W^{(k)} * a^{(k-1)} + b^{k} \tag{2}$$

where $a^{(k)}$ is the activation vector at layer k, $W^{(k)}$ and $b^{(k)}$ are the weight matrix and bias vector, respectively, and f denotes the nonlinear activation function.

The network input is the time series of $I_{Total}(t)$, obtained from the voltage-clamp protocol. The network outputs a vector of predicted biophysical parameters, $\hat{\mathbf{p}} \in \mathbb{R}^{+}$, representing ionic conductances and gating kinetics.

Training and loss function

Networks were trained using mean squared error (MSE) loss and the Adam optimizer (1×10^{-4}). MSE was used as the function evaluation metric (Eq. 3) for predicted model parameters [56,57].

$$MSE = \frac{1}{N}\sum_{i=1}^{n}\left\|p_i - \hat{p}_i\right\|^2 \tag{3}$$

where N is the number of samples, p_i is the ground-truth parameter vector, and $\hat{p}_i$ is the corresponding predicted parameters (the network output). To illustrate how dataset size influences predictive accuracy, a narrower perturbation range of ±20% was used to measure variability around the baseline parameters (Figure 4). We independently sampled 1,000, 10,000, and 200,000 cells at 310 Kelvin (37 °C) to assess how dataset size impacts prediction performance.

For a comprehensive exploration of model variability, a wider ±40% perturbation range was chosen for the 52 parameters (Figure 5). Additionally, all channel kinetics in the synthetic training set were slowed using a Q_{10} factor of 3 to simulate conditions at 298 K (room temperature). The 1.1M-cell dataset size was chosen to ensure broad exploration of the parameter space. This approach helps capture a broad spectrum of experimental differences while ensuring the ionic current properties and action potential behavior stay within realistic limits. Since the full model involved changing many parameters at once, we used a wider range to ensure a diverse enough population, but not so large that the solutions would become mostly non-physiological. Due to the broad parameter variation (±40%), some simulated cells exhibited non-physiological behavior, including non-excitable and non-repolarizing cells. These cells were excluded from the dataset, and only models demonstrating physiologically relevant electrophysiological activity and numerically stable behavior were retained for analysis.

We split the data into training, validation, and test sets using a 70:10:20 ratio, and implemented the network architecture using Keras [58]. The dataset was also divided into 1% sample batches, requiring 70 iterations per epoch. As we monitored the evaluation metrics using the MSE, we continually optimized the voltage clamp protocol to improve the error output and, therefore, get continuously closer to an optimal data generation algorithm for parameter extraction. To prevent overfitting, we computed evaluation metrics using validation data at each training iteration and compared them to values from the training data. When model performance on the training data began to degrade relative to the validation dataset, tuning of the network hyperparameters stopped. Model accuracy was assessed using median mean absolute error (MAE) across parameters. This approach allowed us to extract key ion channel parameters from whole-cell currents, ensuring that the reconstructed ionic current and AP traces closely mirror the physiological behavior of individual cardiac cells.

Population generation and drug testing

We generated a population by applying ± 20% perturbations to all parameters associated with six major ionic currents (I_{K1}, I_{Kr}, I_{Ks}, I_{CaL}, I_{Na}, I_f) around the network-predicted parameter sets for Cell 1 and Cell 2, both originating from the human iPSC lines iPS 6-9-9T.B. This procedure produced a 4000-cell population capturing within-line variability. For the EAD population, we used a narrower perturbation range (±20%) to focus on the parameter space near the baseline model. This allowed us to examine EAD and highlight parameter combinations that promote EAD formation, which only required (±20%) perturbation. The cell population captures intrinsic cell-to-cell variability within a shared genetic background. Simulations were performed under control conditions and in the presence of E-4031 concentrations using a Hill-type hERG block model.

Simulation of I_{Kr} Blockade

To simulate the inhibitory effects of E-4031 on I_{Kr} current, we decreased the peak conductance, G_{Kr} of each of these independent channels in a concentration-dependent fashion using a concentration response relationship with a Hill coefficient of 1 (n = 1) as follows:

$$G_{Kr} = G_{Kr,\max} * \left(\frac{1}{1+([E4031]/IC_{50})^{n}} \right)$$

where $G_{Kr,max}$ is the nominal conductance value obtained from each ventricular myocyte model, and the IC_{50} (7 nM) [44,59] is the concentration of drug that produces a 50% inhibition of the targeted transmembrane current or intracellular organelle.

Simulation and analysis code was written in C++, Python, TensorFlow and MATLAB R2018a (The MathWorks, Natick, MA, USA). The code was executed on two systems: a BIZON ZX5500 workstation with a 2.7 GHz 64-core AMD Threadripper Pro processor, an NVIDIA H100 GPU, and Python version 3.12, TensorFlow version 2.19; and a Mercury GPU408 server with a 2.55 GHz 192-core AMD EPYC 9684 processor, four NVIDIA H100 GPUs, and Python version 3.10, TensorFlow version 2.18. Compilation was performed using GCC version 10.5. Numerical results were visualized in MATLAB R2018a.

Supplementary Information

iPSC-CM electrophysiological model gating kinetics parameters

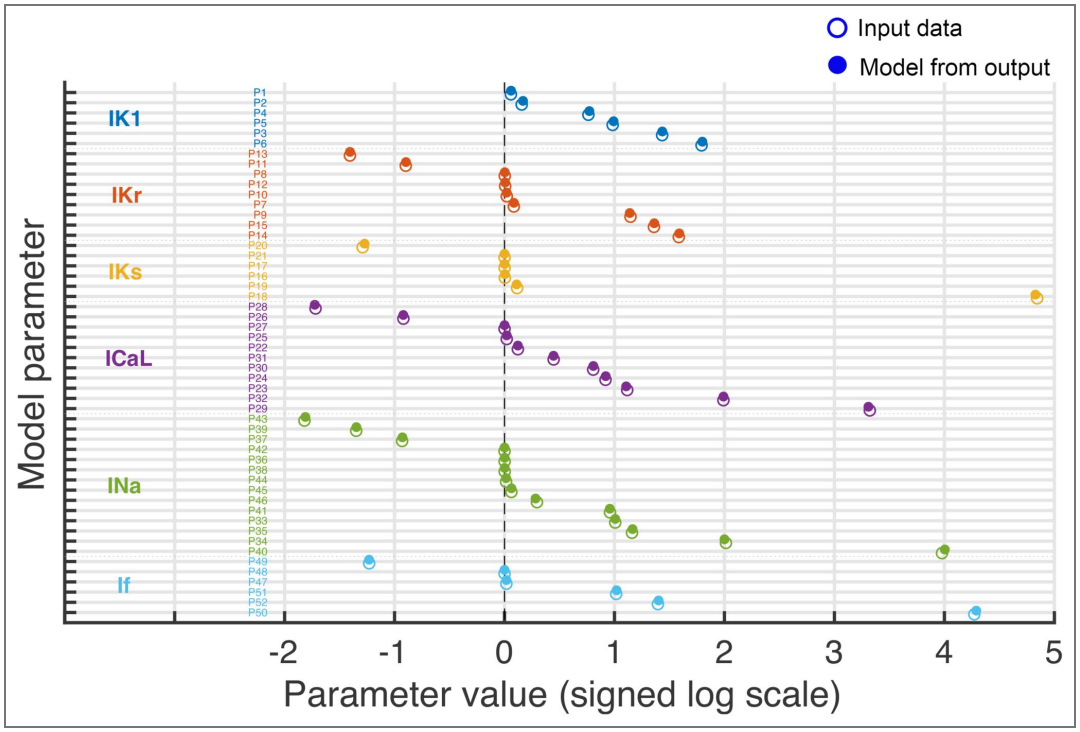


Figure S1. Comparison of model parameter values before and after perturbation. Each parameter (see Table 1 below) is represented by two markers: open circles denote baseline values and filled circles denote updated values. Parameters are grouped by ionic current (IK1, IKr, IKs, ICaL, INa, If), with colors indicating group membership. Within each group, parameters are ordered by the transformed value of the updated condition. Values are displayed on a signed logarithmic scale ($\text{sign}(x)\cdot\log_{10}(1+|x|)$) to accommodate both positive and negative values across a wide dynamic range. The vertical dashed line indicates zero. Horizontal dotted lines separate parameter groups.

IK1		IKr		IKs		ICaL		INa		If	
P1	G_{K1}	P7	GKr	P16	GKs	P22	p_{CaL}	P33	G_{Na}	P47	G_f
P2	xK_{11}	P8	Xr_{1_1}	P17	ks_1	P23	d_1	P34	m_1	P48	xF_1
P3	XK_{12}	P9	Xr_{1_2}	P18	ks_2	P24	d_2	P35	m_2	P49	xF_2
P4	XK_{13}	P10	Xr_{1_5}	P19	ks_5	P25	d_5	P36	m_5	P50	xF_5
P5	XK_{14}	P11	Xr_{1_6}	P20	ks_6	P26	d_6	P37	m_6	P51	xF_6
P6	XK_{15}	P12	Xr_{2_1}	P21	τ_{ks_const}	P27	f_1	P38	h_1	P52	xF_{const}
		P13	Xr_{2_2}			P28	f_2	P39	h_2		
		P14	Xr_{2_5}			P29	f_5	P40	h_5		
		P15	Xr_{2_6}			P30	f_6	P41	h_6		
						P31	τ_{d_const}	P42	j_1		
						P32	τ_{f_const}	P43	j_2		
								P44	τ_{m_const}		
								P45	τ_{h_const}		
								P46	τ_{j_const}		

Table 1. Definitions and correspondence of model parameters (P1–P52) with the variables used in the mathematical equations below.

iPSC-CM electrophysiological model equations

The 52 biophysical variables are highlighted in red in the equations below.

Inward rectifier potassium current (I_{K1})

$$\alpha_{xK1} = xK_{11} \times \exp\left(\frac{V + xK_{13}}{xK_{12}}\right)$$

$$\beta_{xK1} = \exp\left(\frac{V + xK_{15}}{xK_{14}}\right)$$

$$I_{K1} = G_{K1} \times \sqrt{\frac{K_O}{5.4}} \times x_{act,\infty} \times (V - E_K)$$

Rapid delayed rectifier potassium current (I_{Kr})

$$\alpha_{Xr1} = Xr_{1\_1} \times \exp\left(\frac{V}{Xr_{1\_2}}\right)$$

$$\beta_{Xr1} = Xr_{1\_3} \times \exp\left(\frac{V}{Xr_{1\_4}}\right)$$

$$Xr_{1\_3} = Xr_{1\_5} \times Xr_{1\_1}$$

$$Xr_{1\_4} = \frac{1}{\frac{1}{Xr_{1\_2}} + \frac{1}{Xr_{1\_6}}}$$

$$\alpha_{Xr2} = Xr_{2\_1} \times \exp\left(\frac{V}{Xr_{2\_2}}\right)$$

$$\beta_{Xr2} = Xr_{2\_3} \times \exp\left(\frac{V}{Xr_{2\_4}}\right)$$

$$Xr_{2\_3} = Xr_{2\_5} \times Xr_{2\_1}$$

$$Xr_{2\_4} = \frac{1}{\frac{1}{Xr_{2\_2}} + \frac{1}{Xr_{2\_6}}}$$

$$I_{Kr} = G_{Kr} \times \sqrt{\frac{K_O}{5.4}} \times x_{act} \times x_{inact} \times (V - E_K)$$

Slow delayed rectifier potassium current (I_{Ks})

$$\alpha_{Xs} = ks_1 \times \exp\left(\frac{V}{ks_2}\right)$$
$$\beta_{Xs} = ks_3 \times \exp\left(\frac{V}{ks_4}\right)$$
$$\tau_{Xs} = \frac{1}{\alpha_{Xs} + \beta_{Xs}} + \tau_{ks\_const}$$

$$I_{Ks} = G_{Ks} \times x_{act}^2 \times (V - E_K)$$

L-type Ca^{2+} current (I_{CaL})

$$\alpha_d = d_1 \times \exp\left(\frac{V}{d_2}\right)$$
$$\beta_d = d_3 \times \exp\left(\frac{V}{d_4}\right)$$
$$d_3 = d_5 \times d_1$$
$$d_4 = \frac{1}{\frac{1}{d_2} + \frac{1}{d_6}}$$
$$\tau_d = \frac{1}{\alpha_d + \beta_d} + \tau_{d\_const}$$

$$\alpha_f = f_1 \times \exp\left(\frac{V}{f_2}\right)$$
$$\beta_f = f_3 \times \exp\left(\frac{V}{f_4}\right)$$
$$f_3 = f_5 \times f_1$$
$$f_4 = \frac{1}{\frac{1}{f_2} + \frac{1}{f_6}}$$
$$\tau_f = \frac{1}{\alpha_f + \beta_f} + \tau_{f\_const}$$

$$I_{CaL} = P_{CaL,y} \times x_{act} \times x_{inact} \times x_{inact,Ca} \times z_y^2 \times \frac{VF^2}{RT} \times \gamma_y \times \frac{[y]_i \times \exp(z_y VF/RT) - [y]_0}{\exp(z_y VF/RT) - 1}$$

Where y is Ca^{2+}, Na^+, or K^+, $P_{B)C}$ is the permeability ion y, R is the gas constant, z_y is the valence of ion y, and γ_y is to activity coefficient for ion y.

Sodium Current (I_{Na})

$$\alpha_m = m_1 \times \exp\left(\frac{V}{m_2}\right)$$
$$\beta_m = m_3 \times \exp\left(\frac{V}{m_4}\right)$$
$$m_3 = m_5 \times m_1$$
$$m_4 = \frac{1}{\frac{1}{m_2} + \frac{1}{m_6}}$$
$$\tau_m = \frac{1}{\alpha_m + \beta_m} + \tau_{m\_const}$$

$$\alpha_h = h_1 \times \exp\left(\frac{V}{h_2}\right)$$
$$\beta_h = h_3 \times \exp\left(\frac{V}{h_4}\right)$$
$$h_3 = h_5 \times h_1$$
$$h_4 = \frac{1}{\frac{1}{h_2} + \frac{1}{h_6}}$$
$$\tau_h = \frac{1}{\alpha_h + \beta_h} + \tau_{h\_const}$$

$$\alpha_j = j_1 \times \exp\left(\frac{V}{j_2}\right)$$
$$\beta_j = j_3 \times \exp\left(\frac{V}{j_4}\right)$$
$$\tau_j = \frac{1}{\alpha_j + \beta_j} + \tau_{j\_const}$$

$$j_5 = h_5$$
$$j_6 = h_6$$
$$j_3 = j_5 \times j_1$$
$$j_4 = \frac{1}{\frac{1}{j_2} + \frac{1}{j_6}}$$

$$I_{Na} = G_{Na} \times m^3 \times h \times j \times (V - E_{Na})$$

Funny/HCN current (I_f)

$$\alpha_{Xf} = xF_1 \times \exp\left(\frac{V}{xF_2}\right)$$

$$\beta_{Xf} = xF_3 \times \exp\left(\frac{V}{xF_4}\right)$$
$$xF_3 = xF_5 \times xF_1$$

$$xF_4 = \frac{1}{\frac{1}{xF_2} + \frac{1}{xF_6}}$$
$$\tau_{Xf} = \frac{1}{\alpha_{Xf} + \beta_{Xf}} + xF_{const}$$

$$I_f = G_f \times x_{act} \times (V - E_f)$$

For detailed model equations and parameters, please refer to the Kernik-Clancy model[44].

Data availability

Due to the large size of the synthetic datasets (~1.1M cells), the raw data are not hosted online. However, a representative dataset of 100 cells is provided as an example, and the complete codebase for dataset generation, along with parameter distributions and simulation protocols, is

publicly available at GitHub (*https://github.com/ClancyLabUCD/Digital-Twin-for-the-Win-Personalized-Cardiac-Electrophysiology*). This allows full regeneration of the training and validation datasets as described in the study.

Additional information

Sources of funding

This work was supported in part by the National Institutes of Health under grants R01HL128537, OT2OD026580, and R01HL17400 (to C. E. Clancy and L. F. Santana), and R01HL159492 (to D. K. Lieu). G. H-Hernandez was supported by the UC Davis Chancellor's Postdoctoral Fellowship. Additional support from The UC Davis Center for Precision Medicine and Data Sciences was provided to C. E. Clancy, P.-C. Yang and computing facilities.

Contributions

PCY, MTJ, and CC designed the research study, conducted simulations, acquired data, and analyzed data. DKL and RS sample preparation and differentiation of the iPSC lines. G.H.H provided experimental support. LFS conducted the electrophysiological recordings. PCY, LFS and CEC analyzed data, and wrote the manuscript.

Funding

Funder	Grant reference number	Author
HHS | NIH | National Heart, Lung, and Blood Institute (NHLBI)	R01HL128537	L Fernando Santana Colleen E Clancy
HHS | NIH | National Heart, Lung, and Blood Institute (NHLBI)	OT2OD026580	L Fernando Santana Colleen E Clancy
HHS | NIH | National Heart, Lung, and Blood Institute (NHLBI)	R01HL17400	L Fernando Santana Colleen E Clancy
HHS | NIH | National Heart, Lung, and Blood Institute (NHLBI)	R01HL159492	Deborah K Lieu
UC | University of California, Davis (UCD)	Chancellor's Postdoctoral Fellowship	Gonzalo Hernandez-Hernandez
UC | University of California, Davis (UCD)	Center for Precision Medicine and Data Sciences	Pei-Chi Yang Colleen E Clancy

Author ORCID iDs

Pei-Chi Yang: https://orcid.org/0000-0002-5753-1131
Colleen E Clancy: https://orcid.org/0000-0001-6849-4885

Additional files

Supplementary Information.

References

1 **Serrano R,** et al. (2023) A deep learning platform to assess drug proarrhythmia risk. *Cell Stem Cell* **30**:86-95.e84, https://doi.org/10.1016/j.stem.2022.12.002 | PubMed

2 **Gintant G,** Sager PT, Stockbridge N (2016) Evolution of strategies to improve preclinical cardiac safety testing. *Nat Rev Drug Discov* **15**:457-471 https://doi.org/10.1038/nrd.2015.34 | PubMed

3 **Laverty H,** et al. (2011) How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines?. *Br J Pharmacol* **163**:675-693 https://doi.org/10.1111/j.1476-5381.2011.01255.x | PubMed

4 **Akhtar A** (2015) The flaws and human harms of animal experimentation. *Camb Q Healthc Ethics* **24**:407-419 https://doi.org/10.1017/S0963180115000079 | PubMed

5 **Bracken MB** (2009) Why animal studies are often poor predictors of human reactions to exposure. *J R Soc Med* **102**:120-122 https://doi.org/10.1258/jrsm.2008.08k033 | PubMed

6 **Franco R,** Cedazo-Minguez A (2014) Successful therapies for Alzheimer's disease: why so many in animal models and none in humans?. *Front Pharmacol* **5** https://doi.org/10.3389/fphar.2014.00146 | PubMed

7 **Heist EK,** Ruskin JN (2010) Drug-induced arrhythmia. *Circulation* **122**:1426-1435 https://doi.org/10.1161/CIRCULATIONAHA.109.894725 | PubMed

8 **Sorger PK,** Allerheiligen SRB (2011) Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms. *Tech. Rep*

9 **Fink M,** et al. (2011) Cardiac cell modelling: observations from the heart of the cardiac physiome project. *Prog Biophys Mol Biol* **104**:2-21 https://doi.org/10.1016/j.pbiomolbio.2010.03.002 | PubMed

10 **Passini E,** et al. (2017) Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. *Front Physiol* **8** https://doi.org/10.3389/fphys.2017.00668 | PubMed

11 **Tveito A,** Jaeger KH, Maleckar MM, Giles WR, Wall S (2020) Computational translation of drug effects from animal experiments to human ventricular myocytes. *Sci Rep* **10** https://doi.org/10.1038/s41598-020-66910-0 | PubMed

12 **Tveito A,** et al. (2018) Inversion and computational maturation of drug response using human stem cell derived cardiomyocytes in microphysiological systems. *Sci Rep* **8** https://doi.org/10.1038/s41598-018-35858-7 | PubMed

13 **Lieu DK,** et al. (2013) Mechanism-based facilitated maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ Arrhythm Electrophysiol* **6**:191-201 https://doi.org/10.1161/CIRCEP.111.973420 | PubMed

14 **De La Mata A,** et al. (2019) BIN1 Induces the Formation of T-Tubules and Adult-Like Ca(2+) Release Units in Developing Cardiomyocytes. *Stem Cells* **37**:54-64 https://doi.org/10.1002/stem.2927 | PubMed

15 **Thomas D,** Shenoy S, Sayed N (2021) Building Multi-Dimensional Induced Pluripotent Stem Cells-Based Model Platforms to Assess Cardiotoxicity in Cancer Therapies. *Front Pharmacol* **12** https://doi.org/10.3389/fphar.2021.607364 | PubMed

16 **de Korte T,** et al. (2020) Unlocking Personalized Biomedicine and Drug Discovery with Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes: Fit for Purpose or Forever Elusive?. *Annu Rev Pharmacol Toxicol* **60**:529-551 https://doi.org/10.1146/annurev-pharmtox-010919-023309 | PubMed

17 **Hnatiuk AP,** Briganti F, Staudt DW, Mercola M (2021) Human iPSC modeling of heart disease for drug development. *Cell Chem Biol* **28**:271-282 https://doi.org/10.1016/j.chembiol.2021.02.016 | PubMed

18 **Sayed N,** Liu C, Wu JC (2016) Translation of Human-Induced Pluripotent Stem Cells: From Clinical Trial in a Dish to Precision Medicine. *J Am Coll Cardiol* **67**:2161-2176 https://doi.org/10.1016/j.jacc.2016.01.083 | PubMed

19 **Burridge PW,** et al. (2016) Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity. *Nat Med* **22**:547-556 https://doi.org/10.1038/nm.4087 | PubMed

20 **Bizy A,** Klos M (2020) Optimizing the Use of iPSC-CMs for Cardiac Regeneration in Animal Models. *Animals* **10** https://doi.org/10.3390/ani10091561 | PubMed

21 **Wu P,** et al. (2021) Maturation strategies and limitations of induced pluripotent stem cell-derived cardiomyocytes. *Biosci Rep* **41** https://doi.org/10.1042/BSR20200833 | PubMed

22 **Vuckovic S,** et al. (2022) Characterization of cardiac metabolism in iPSC-derived cardiomyocytes: lessons from maturation and disease modeling. *Stem Cell Res Ther* **13** https://doi.org/10.1186/s13287-022-03021-9 | PubMed

23 **Hwang HS,** et al. (2015) Comparable calcium handling of human iPSC-derived cardiomyocytes generated by multiple laboratories. *J Mol Cell Cardiol* **85**:79-88 https://doi.org/10.1016/j.yjmcc.2015.05.003 | PubMed

24 **Veerman CC,** et al. (2015) Immaturity of human stem-cell-derived cardiomyocytes in culture: fatal flaw or soluble problem?. *Stem Cells Dev* **24**:1035-1052 https://doi.org/10.1089/scd.2014.0533 | PubMed

25 **Fink M,** Noble D (2009) Markov models for ion channels: versatility versus identifiability and speed. *Philos Trans A Math Phys Eng Sci* **367**:2161-2179 https://doi.org/10.1098/rsta.2008.0301 | PubMed

26 **Groenendaal W,** et al. (2015) Cell-specific cardiac electrophysiology models. *PLoS Comput Biol* **11**:e1004242 https://doi.org/10.1371/journal.pcbi.1004242 | PubMed

27 **Moreno JD,** Lewis TJ, Clancy CE (2016) Parameterization for In-Silico Modeling of Ion Channel Interactions with Drugs. *PLoS One* **11**:e0150761 https://doi.org/10.1371/journal.pone.0150761 | PubMed

28 **Guo T,** Al Abed A, Lovell NH, Dokos S (2013) Optimisation of a generic ionic model of cardiac myocyte electrical activity. *Comput Math Methods Med* **2013**:706195 https://doi.org/10.1155/2013/706195 | PubMed

29 **Nelder J,** Mead R (1965) A simplex method for function minimization. *Computer Journal* **7**:308-313 https://doi.org/10.1093/COMJNL/7.4.308

30 **Dokos S,** Lovell NH (2004) Parameter estimation in cardiac ionic models. *Prog Biophys Mol Biol* **85**:407-431 https://doi.org/10.1016/j.pbiomolbio.2004.02.002 | PubMed

31 **Hui BB,** Dokos S, Lovell NH (2007) Parameter identifiability of cardiac ionic models using a novel CellML least squares optimization tool. *Annu Int Conf IEEE Eng Med Biol Soc* **2007**:5307-5310 https://doi.org/10.1109/IEMBS.2007.4353539 | PubMed

32 **Lei CL,** et al. (2020) Accounting for variability in ion current recordings using a mathematical model of artefacts in voltage-clamp experiments. *Philos Trans A Math Phys Eng Sci* **378** https://doi.org/10.1098/rsta.2019.0348 | PubMed

33 **Clark AP,** Wei S, Fullerton K, Krogh-Madsen T, Christini DJ (2024) Single-cell ionic current phenotyping explains stem cell-derived cardiomyocyte action potential morphology. *Am J Physiol Heart Circ Physiol* **326**:H1146-H1154 https://doi.org/10.1152/ajpheart.00063.2024 | PubMed

34 **Kirsch GE,** et al. (2004) Variability in the measurement of hERG potassium channel inhibition: effects of temperature and stimulus pattern. *J Pharmacol Toxicol Methods* **50**:93-101 https://doi.org/10.1016/j.vascn.2004.06.003 | PubMed

35 **Yao JA,** et al. (2005) Estimation of potency of HERG channel blockers: impact of voltage protocol and temperature. *J Pharmacol Toxicol Methods* **52**:146-153 https://doi.org/10.1016/j.vascn.2005.04.008 | PubMed

36 **Milnes JT,** Witchel HJ, Leaney JL, Leishman DJ, Hancox JC (2010) Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 degrees C: Cisapride versus dofetilide. *J Pharmacol Toxicol Methods* **61**:178-191 https://doi.org/10.1016/j.vascn.2010.02.007 | PubMed

37 **Narsinh KH,** et al. (2011) Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J Clin Invest* **121**:1217-1221 https://doi.org/10.1172/JCI44635 | PubMed

38 **Doss MX,** et al. (2012) Maximum diastolic potential of human induced pluripotent stem cell-derived cardiomyocytes depends critically on I(Kr). *PLoS One* **7**:e40288 https://doi.org/10.1371/journal.pone.0040288 | PubMed

39 **Du DT,** Hellen N, Kane C, Terracciano CM (2015) Action potential morphology of human induced pluripotent stem cell-derived cardiomyocytes does not predict cardiac chamber specificity and is dependent on cell density. *Biophys J* **108**:1-4 https://doi.org/10.1016/j.bpj.2014.11.008 | PubMed

40 **Britton OJ,** et al. (2013) Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proc Natl Acad Sci U S A* **110**:E2098-2105 https://doi.org/10.1073/pnas.1304382110 | PubMed

41 **Gong JQX,** Sobie EA (2018) Population-based mechanistic modeling allows for quantitative predictions of drug responses across cell types. *NPJ Syst Biol Appl* **4** https://doi.org/10.1038/s41540-018-0047-2 | PubMed

42 **Ni H,** Morotti S, Grandi E (2018) A Heart for Diversity: Simulating Variability in Cardiac Arrhythmia Research. *Front Physiol* **9** https://doi.org/10.3389/fphys.2018.00958 | PubMed

43 **Muszkiewicz A,** et al. (2016) Variability in cardiac electrophysiology: Using experimentally-calibrated populations of models to move beyond the single virtual physiological human paradigm. *Prog Biophys Mol Biol* **120**:115-127 https://doi.org/10.1016/j.pbiomolbio.2015.12.002 | PubMed

44 **Kernik DC,** et al. (2019) A computational model of induced pluripotent stem-cell derived cardiomyocytes incorporating experimental variability from multiple data sources. *J Physiol* **597**:4533-4564 https://doi.org/10.1113/JP277724 | PubMed

45 **Ma D,** et al. (2015) Characterization of a novel KCNQ1 mutation for type 1 long QT syndrome and assessment of the therapeutic potential of a novel IKs activator using patient-specific induced pluripotent stem cell-derived cardiomyocytes. *Stem Cell Res Ther* **6** https://doi.org/10.1186/s13287-015-0027-z | PubMed

46 **Ma J,** et al. (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**:H2006-2017 https://doi.org/10.1152/ajpheart.00694.2011 | PubMed

47 **Herron TJ,** et al. (2016) Extracellular Matrix-Mediated Maturation of Human Pluripotent Stem Cell-Derived Cardiac Monolayer Structure and Electrophysiological Function. *Circ Arrhythm Electrophysiol* **9**:e003638 https://doi.org/10.1161/CIRCEP.113.003638 | PubMed

48 **Li M,** et al. (2017) Overexpression of KCNJ2 in induced pluripotent stem cell-derived cardiomyocytes for the assessment of QT-prolonging drugs. *J Pharmacol Sci* **134**:75-85 https://doi.org/10.1016/j.jphs.2017.05.004 | PubMed

49 **Es-Salah-Lamoureux Z,** et al. (2016) HIV-Tat induces a decrease in I(Kr) and I(Ks)via reduction in phosphatidylinositol-(4,5)-bisphosphate availability. *J Mol Cell Cardiol* **99**:1-13 https://doi.org/10.1016/j.yjmcc.2016.08.022 | PubMed

50 **Veerman CC,** et al. (2016) hiPSC-derived cardiomyocytes from Brugada Syndrome patients without identified mutations do not exhibit clear cellular electrophysiological abnormalities. *Sci Rep* **6** https://doi.org/10.1038/srep30967 | PubMed

51 **Bellin M,** et al. (2013) Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. *EMBO J* **32**:3161-3175 https://doi.org/10.1038/emboj.2013.240 | PubMed

52 **Garg P,** et al. (2018) Genome Editing of Induced Pluripotent Stem Cells to Decipher Cardiac Channelopathy Variant. *J Am Coll Cardiol* **72**:62-75 https://doi.org/10.1016/j.jacc.2018.04.041 | PubMed

53 **DeBoever C,** et al. (2017) Large-Scale Profiling Reveals the Influence of Genetic Variation on Gene Expression in Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **20**:533-546.e537, https://doi.org/10.1016/j.stem.2017.03.009 | PubMed

54 **Krogh A** (2008) What are artificial neural networks?. *Nature biotechnology* **26**:195-197 https://doi.org/10.1038/nbt1386 | PubMed

55 **Carugo O,** Eisenhaber F (2010) *Data mining techniques for the life sciences* Springer.

56 **Goodfellow I,** Bengio Y, Courville A (2016) *Deep learning* MIT press.

57 **Murphy KP** (2012) *Machine learning: a probabilistic perspective* MIT press.

58 **Prakash KB,** Kannan R, Alexander SA, Kanagachidambaresan GR (2021) *Advanced Deep Learning for Engineers and Scientists: A Practical Approach* Springer.

59 **Harris K,** et al. (2013) Comparison of electrophysiological data from human-induced pluripotent stem cell-derived cardiomyocytes to functional preclinical safety assays. *Toxicol Sci* **134**:412-426 https://doi.org/10.1093/toxsci/kft113 | PubMed

Peer reviews

Reviewer #2 (Public review):

Summary:

The authors present a computational framework for generating "cell-specific" digital twins of human iPSC-CMs from a single optimized voltage clamp recording. Using deep learning trained on > 1 million artificial cells, the authors demonstrate that the model can infer 52 biophysical parameters governing 6 major ionic currents, and the resulting digital twins can reproduce experimentally recorded action potentials.

Comments on revised version:

The authors propose an interesting platform for digital twin construction of iPSC-CMs using an AI-based approach. However, regarding the fundamental concerns raised in the previous review round "lack of experimental validation" and "overstatement of the claims", the authors have merely added text to the "Limitations" in the Discussion, without providing any new wet-lab experimental data. This cosmetic revision fails to demonstrate the scientific validity of the platform, and the core issues remain completely unresolved.

I think the authors need to either provide substantial additional experimental data or drastically tone down the claims throughout the manuscript based on the following three major concerns.

(1) Lack of wet validation

The authors show that their AI model can infer 52 parameters from a single patch-clamp recording and reproduce the overall action potential waveform. However, the most critical validation (whether the individual ion channel parameters, such as IKr/ICaL, inferred by the AI actually match the true parameters of that specific cell) is still missing. Without a direct head-to-head comparison between the parameters inferred by the model and the exact values measured using conventional wet experiments, it is impossible to determine whether the platform is providing accurate prediction (or merely performing a curve-fitting).

(2) Absence of experimental validation for drug response simulations (Cell 1 vs. Cell 2)

In Figure 6, the authors present a simulation result where the administration of an IKr blocker (E-4031) induces EADs in the digital twin of Cell 1, but not Cell 2. However, there is absolutely no wet-lab validation for this prediction. Unless the authors actually administer the same drug to the live Cell 1 and Cell 2 from which the recordings were taken, this "computational drug response prediction" remains purely hypothetical. There is no evidence provided that the prediction accurately reflects real biological responses.

(3) Significant overstatement regarding "inter-individual variability" and "personalized medicine"

The authors state in the very first sentence of the Abstract: "Individual variability shapes how diseases manifest, how patients respond to therapy, and how rare phenotypes arise". However, this opening sentence is severely disconnected from the actual conclusions and data presented in this study. The platform can capture only "cell-to-cell variability within the same dish" (which is not even validated), and thus claiming "patient-to-patient differences" is an overstatement.

https://doi.org/10.7554/eLife.110013.2.sa2

Reviewer #3 (Public review):

Summary:

This work use convolution neural network to optimize a voltage clamp protocol to identify features and parameters from human pluripotent stem cell-derived cardiomyocytes.

Strengths:

The major strength is the methodology used to bridge in silico prediction of cell behavior and mechanistic insights from experimental dataset.

Comments on revised version.

As highlighted by the authors, due to the variability of the hPSC-CM model, to increase the applicability of this method, additional experimental dataset from different hPSC-CM lines would increase the translation of this approach.

I personally found that the detailed description of the methods, including the rationale of including/excluding some parameters, is extremely helpful to whoever would like to use this approach in their research.

https://doi.org/10.7554/eLife.110013.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *This study presents an interesting approach for finding electrophysiological models that match experimental patch-clamp data. The authors develop a new method for deriving optimized current clamp protocols by training a neural network on synthetic data. This optimized current clamp is then used on both computational training data and on experimental data to predict current gating and conductance parameters that correctly reconstruct the electrical phenotype.*
>
> *Strengths:*
>
> *(1) The fitting of gating variables through an optimized patch clamp protocol is interesting.*
>
> *(2) The inclusion of experimental data is important, and the approach is shown to be effective in fitting them.*
>
> *Weaknesses:*
>
> *(1) Some clarity is necessary on the generation and selection of variable IPSC models. With such a large variation in so many parameters, I would expect some resulting parameters to generate non-realistic phenotypes, quiescent cells, etc. Are all 200,000 or 1,100,000 generated cells viable? Or are they selected somehow for realistic cell properties?*

Thank you for this important point. We agree that broad parameter variation can generate non-physiological model behavior. Indeed, with the +/-40% perturbation range, some simulated cells produced non-realistic outputs, including quiescent behavior, and failure to generate a complete action potential. These cases were excluded from the dataset. As a result, only cells exhibiting physiologically meaningful and numerically stable behavior were retained for further analysis. We have clarified this selection procedure in the Methods section. We applied a large variation to ensure that all possible combinations and morphologies were included in the training and testing data so the model would readily ingest new data and perform robustly.

> *(2) The error shown in Figure 4 between different population sizes is not completely explained in the text - there seems to be a minimal difference between a population of 1,000 and 10,000, followed by a very good fit at 200,000. Is there a particular threshold that needs to be crossed where the error drops off? Related, how was the 200,000 number chosen?*

Thank you for this observation. We agree that the decrease in error shows a gradual performance improvement as the population size increases, rather than a strict cutoff. As shown in Figure 4, the difference between 1,000 and 10,000 samples is small, but as we continue to increase and get to around 200,000 samples, we see strong error minimization. This indicates how much training data is needed for optimal model performance. This improvement is due to better coverage of the high-dimensional parameter space, which helps the network learn the nonlinear relationships between the parameters and outputs.

We tested a range of training data sets and found that above 200,000 training data sets, the model consistently produced low, stable errors and good test-training agreement. The test error decreased with the training error as the population size increased, indicating better generalization and suggesting that the model accurately predicts unseen data rather than overfitting to the training set.

> *(3) Related to the point above, the 1,100,000 population for fitting experimental data also needs a more complete explanation: how was this number chosen, and how does the error compare with the other population sizes shown in Figure 4?*

Thank you for this question. We found that at a training data set size of 1,100,000 we were able to cover the large parameter space induced by +/-40% parameter perturbation. iPSC-CM measurements are known to exhibit high variability, and we wanted to capture the full range in the training data set so the model could ingest a wide range of experimental data. It is trivial to generate new training data, for example, to capture different experimental conditions like temperature differences, mutations, drugs, or ionic variability. We view this flexibility as a substantial strength of the approach. But the large perturbations we show in this study (+/-40%) allow the generation of a very broad range of cellular phenotypes while maintaining physiologically realistic ionic current properties and action potential behavior. Consistent with Figure 4, increasing population size reduces prediction error and improves generalization. The larger dataset provided more stable, accurate predictions when fitting experimental data, without evidence of overfitting.

> *(4) Why are the optimized current clamp protocols different between panels A and B in Figure 5? Are they somehow informed by experimental data?*

Thank you for this question. The stimulation protocol used in panels A and B is identical. Panels A and B show whole-cell currents recorded under the same stimulation conditions as in Figure 3. The differences reflect variability in the underlying whole-cell ionic currents of the model cells rather than differences in the applied protocol. This is exactly the idea: the exact same protocol will generate different whole-cell currents in individual cells, but the model can find parameter sets for all of them.

> *(5) Figure 6D: Is the EAD risk in panel D specific to cell 1, 2, or the pooled variants of both?*

Thank you for this question. We have clarified this point in the revised manuscript. The EAD risk shown in panel D is computed from the pooled variants of both Cell 1 and Cell 2, rather than being specific to either cell individually.

> *(6) How sensitive is the fitting to minor parameter variation? Further, if one were to pick, let's say, the next-best-fitting value, would that fall close to the best one? Is the solution found unique, or are there multiple sets with good fits?*

Traditional optimization methods, such as Nelder–Mead, directly fit the model to the observed data by iteratively minimizing the error for each dataset. As a result, the solution can depend on the initial parameter guess and may converge to different local minima. In contrast, our approach trains a deep learning model on synthetic data generated from the baseline model, learning a mapping from whole-cell currents to the corresponding 52-parameter sets by minimizing prediction error. The mean squared error (MSE) decreases from approximately 10^{-2} to below 10^{-3}, with training and test errors overlapping closely, indicating stable training, good generalization, and accurate reproduction of the observed signals.

The model achieves very low MSE and reproduces the electrophysiological outputs with high fidelity. However, accurate reproduction of the outputs does not imply a unique parameter solution. This is illustrated in Figure S1, where baseline and predicted parameter values show close agreement overall, yet small deviations persist across parameters. This indicates that different parameter combinations can yield similar whole-cell behaviors due to parameter correlations and compensatory effects. In such cases, the model learns to predict a representative parameter set that is most consistent with the training data and loss function, rather than converging to a single unique solution within a fixed numerical tolerance.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *The authors present a computational framework for generating "cell-specific" digital twins of human iPSC-CMs from a single optimized voltage clamp recording. Using deep learning trained on > 1 million artificial cells, the authors demonstrate that the model can infer 52 biophysical parameters governing 6 major ionic currents, and the resulting digital twins can reproduce experimentally recorded action potentials.*
>
> *Strengths:*
>
> *The framework has clear potential for understanding cellular heterogeneity in iPSC-CMs, predicting individual drug responses, and reducing the experimental burden of multiple patch clamp protocols.*
>
> *Weaknesses:*
>
> *There are several concerns about the validation of the model and its clarity. First, the biological variability being modeled in this manuscript is not defined well. It is unclear*

whether the framework addresses cell-to-cell differences within a single differentiation batch, variability across iPSC lines, or donor-to-donor differences. This ambiguity makes it difficult to interpret what the "digital twin populations" actually represent biologically. Second, the main claim, "the digital twins enable drug testing and arrhythmia prediction that would be impractical experimentally", is not experimentally validated. For example, the E-4031 simulations predict EAD rates, but no direct experimental head-to-head comparison is provided to confirm that these predictions are accurate. Third, technical reproducibility and biological representativeness are not assessed. Single voltage clamp recordings are inherently noisy. Without knowing how much variability comes from the recording process (technical variation) vs true biological differences, it is difficult to judge whether observed "cell-specific" parameter differences are meaningful. In addition, the optimized protocol is claimed to be superior to conventional approaches, but again, no experimental comparison is shown.

The authors should address these concerns, with particular emphasis on clarifying the biological context and providing direct experimental validation. Below are detailed specific points:

(1) Ambiguous definition of iPSC-CM heterogeneity. The authors model "typical iPSC-CM heterogeneity" by varying 52 parameters +/- 40% around a baseline model (Figure 1), generating > 1 million synthetic cells. However, the manuscript does not clearly state what biological variability this model is intended to capture. Is this modeling within-line, cell-to-cell variability (e.g., cells from the same dish or differentiation batch that differ due to stochastic gene expression or maturation state)? Or is this modeling between-line or between-donor variability (e.g., genetic background differences, reprogramming efficiency)? This distinction is critical for interpretation. If the goal is to understand why different cells in the same dish behave differently, then training data should reflect that. If the goal is to compare patient lines or disease models, the framework needs validation across multiple donors or lines.

For example, the experimental validation in Figure 5 uses a single iPSC line (iPS-6-9-9T.B), but how many differentiation batches or dishes were tested, or whether cells came from the same preparation are unclear. Another example is that the wide AP diversity in the training population (Figure 1A) is impressive, but there is no demonstration that real experimental cells actually fall within this assumption range of +/- 40%.

From a biological perspective, iPSC-CMs are known to be highly heterogeneous within lines (maturation state, metabolic differences, epigenetic variation, spatial differences within the same dish, etc) and between lines (different donor/genetic background). Thus, please explicitly state whether the +/- 40% variation is intended to model within-line or between-line heterogeneity, and justify this choice with wet experiment data (or reference to experimental literature on iPSC-CM variability). Please clarify how many dishes, differentiation batches, and time points post-differentiation were used for experimental recordings (Figures 5-6). If the framework is intended to generalize across lines from different donors, please test the model on multiple independent iPSC lines (from different donors).

Thank you for this important and insightful comment. The selected ±40% range was chosen to broadly explore all physiologically plausible electrophysiological behaviors, not to match a specific experimental distribution. Our goal was to cover enough behaviors for the model to learn a reliable mapping between responses and ionic parameters.

We recognize that this approach does not explicitly account for variability between lines or donors. We have a current project focused on extending the framework to include multiple iPSC-CMs from patient donors, but given that the model framework successfully reproduces

such a broad range of cell phenotypes, we feel confident that it will readily apply to different genetic backgrounds from patient-specific cells. This study is underway.

We have updated the manuscript to clarify how the modeled variability is interpreted and added a discussion of these limitations. Furthermore, we clarified the experimental conditions, such as the number of differentiation batches and recording settings, in the revised Methods section.

> *(2) Biological representativeness of single-cell measurements.*
>
> *The framework generates digital twins from single voltage clamp recordings. The patch clamp recordings in iPSC-CMs are subject to substantial technical variability. The manuscript does not address a fundamental question: "How representative are the measurements from a single cell on the dish (or line)?" In other words, if I measure one cell from a dish of a million cells, does that cell's digital twin tell me something about the dish as a whole, or just about that one cell? The manuscript presents Cell 1 and Cell 2 (Figures 5-6) as distinct individuals, but it's unclear whether these differences reflect true biological heterogeneity or simply sampling variability. I think the authors should perform replicate recordings on multiple cells (e.g., > 10 cells) from the same dish (same differentiation batch) and quantify how much the inferred parameters vary, and then compare between lines.*

Thank you for this important comment. We agree that the representativeness of single-cell measurements and the impact of technical variability are important considerations in interpreting the results. In this study, the framework is designed to generate digital twins that reflect the electrophysiological properties of individual recorded cells, rather than to directly represent the behavior of the entire cell population within a dish.

As such, differences observed between Cell 1 and Cell 2 are intended to reflect variability at the single-cell level, which may arise from a combination of biological heterogeneity and experimental variability. We agree that systematic replicate recordings across multiple cells are valuable to quantify the relative contributions of biological and technical variability, and to assess the consistency of inferred parameters. However, this is beyond the scope of the current study. We have added clarification in the manuscript to explicitly state this limitation and to outline this as an important direction for future work.

> *(3) No experimental validation of the main claim that in silico populations can replace wet experiments.*
>
> *The most exciting claim in the manuscript is that digital twins enable drug testing and arrhythmia prediction "at scale" without requiring hundreds of patch clamp experiments. Specifically, the authors show that in silico populations derived from two experimental cells (Figure 6C) predict dose-dependent EAD incidence for the IKr blocker E-4031 (Figure 6D), with ~3% of cells showing EADs at 50 nM.*
>
> *However, this prediction is not validated experimentally. If I actually patch 20-30 real iPSC-CMs and apply 50 nM E-4031, will ~3% of them show EADs, as the model predicts? Without this validation, I think the drug testing framework is purely hypothetical. The model may be internally consistent (e.g., Cell 1's twin behaves differently from Cell 2's twin), but there is no evidence that these in silico populations reflect real biological variability in drug response. Please provide experimental validation that justifies the prediction by digital twins.*

Thank you for this important comment. We agree that experimental validation of population-level drug response will be valuable for establishing the quantitative accuracy of the predicted EAD incidence. The E-4031 simulations are intended as a proof-of-concept illustrating how the framework can identify susceptible subpopulations and quantify relative

proarrhythmic risk *in silico*. We agree that direct comparison with large-scale experimental datasets is a key next step, and we are working hard to get the study funded so that we can perform those experiments and bring this technology to scale.

> *(4) Experimental validation and head-to-head comparison of optimized protocol.*
>
> *The authors claim that their deep learning-optimized voltage clamp protocol (Figure 3, Figure 4A) is superior to conventional approaches, but they have not validated this experimentally by doing a head-to-head comparison. The manuscript does not compare the optimized protocol to any published voltage clamp designs. If the optimized protocol is genuinely easier to implement and more informative than existing approaches, this would be a major practical advance. But without side-by-side comparison, it is impossible to judge whether the optimization made a real difference.*

Thank you for your comment. We agree that comparing directly with traditional voltage-clamp protocols through experiments would be useful. In this study, our main aim was to show that the optimized protocol enhances parameter inference within the modeling framework, not to prove experimental superiority. We have clarified this point in the revised version.

> ***Reviewer #3 (Public review):***
>
> *Summary:*
>
> *This work uses a convolutional neural network to optimize a voltage clamp protocol to identify features and parameters from human pluripotent stem cell-derived cardiomyocytes.*
>
> *Yang et al. introduce an innovative experimental framework that integrates computational modeling and deep learning to generate a digital twin of human pluripotent stem cell-derived cardiomyocytes (hPSC-CMs).*
>
> *Strengths:*
>
> *The major strength is the methodology used to bridge in silico prediction of cell behavior and mechanistic insights from the experimental dataset.*
>
> *The approach used in this study represents a significant step toward precision medicine by enabling in silico prediction of cellular behavior and mechanistic insight from experimental datasets. The study addresses an important and timely challenge in stem cell-based and personalized medicine, and the authors compellingly leverage state-of-the-art methods alongside strong expertise in computational modeling and cardiac electrophysiology*
>
> *Weaknesses:*
>
> *While the overall approach is highly compelling and the potential impact is substantial, there are two areas where clarification and refinement, particularly in the phrasing and framing used throughout the manuscript, would further strengthen the work.*
>
> *(1) While the overall goal of the study is compelling, the manuscript would benefit from clearer articulation of how the proposed framework is intended to be used in practice. In particular, it is not entirely clear whether the authors envision this approach as:*
>
> *(a) a method to extract population-level trends that, when paired with biological data, enhance statistical power and interpretability, or*
>
> *(b) a strategy capable of constructing a population-based model from limited single-cell recordings. If the latter is intended, additional guidance on the number of action*

> *potentials required per cell and the assumptions underlying this extrapolation would greatly clarify the scope and applicability of the method.*

Thank you for this thoughtful comment. We agree that the intended use of the framework should be more clearly articulated. In this study, we generate a large synthetic population of iPSC-CM models by varying 52 biophysical parameters governing key ionic currents. A neural network is trained on simulated whole-cell current responses to learn a mapping between current profiles and model parameters. Experimental recordings are then used as inputs to this trained model to infer ionic parameters, rather than directly fitting the model to data. This enables individual recordings to be interpreted within a large, physiologically plausible parameter space and supports population-level analysis of electrophysiological variability. The primary goal of the framework is therefore to facilitate mechanistic interpretation of variability and relate experimental observations to underlying ionic currents. But the longer-term intended goal is to develop digital twins from patient-derived cell lines and then use populations constructed from patient-specific digital twins to screen therapeutics and identify arrhythmia marker vulnerability in a very thorough and high-throughput way. We have clarified this in the revised manuscript.

> *(2) The manuscript would also benefit from a clearer explanation of how electrophysiological heterogeneity observed in hPSC-CMs is linked to inter-patient variability. Although the authors state that this framework can be generalized to compare patient-specific hiPSC-CM lines, it remains unclear how this generalization is achieved, given the substantial sources of variability intrinsic to hiPSC-CMs (e.g., batch effects, reprogramming strategy, differentiation protocol, and maturation state). As acknowledged by the authors, addressing this level of variability likely requires large datasets; further clarification of how the proposed approach mitigates or accommodates these challenges would strengthen the translational claims.*
>
> *Below are my suggestions that could help strengthen the claims in the manuscript:*
>
> *(1) Adding a dedicated section describing the electrophysiological phenotype of the hPSC-CMs used in this study would help justify the choice of the underlying ionic model and the selection of the six ion currents analyzed. These currents are not only developmentally regulated but may also vary substantially across different hPSC-CM lines, which has implications for generalizability.*

Thank you for this important suggestion. We agree that providing additional context on the electrophysiological phenotype of the hPSC-CMs strengthens the rationale for both the underlying ionic model and the selection of currents analyzed.

We have expanded the Methods section to clarify this point. Briefly, the ionic currents were selected based on the Kernik-Clancy iPSC-CM model developed in our prior work, which was specifically designed to capture the range of electrophysiological variability observed within an iPSC-CM cell line using a population-based framework. In this model, variation in key ionic conductances is sufficient to reproduce the diversity of action potential morphologies, spontaneous activity, and repolarization dynamics commonly reported experimentally, while avoiding non-physiological behaviors.

Accordingly, we focused on six primary ionic currents that are known to play dominant roles in shaping action potential characteristics and variability in iPSC-CMs. This selection reflects a balance between model parsimony and physiological relevance, enabling the framework to capture the expected spectrum of variability within a given cell line. We also note that the framework is extensible, and additional currents or alternative parameterizations can be incorporated to account for differences across cell lines, donors, or experimental conditions in future studies. See updated discussion.

> *(2) If feasible, inclusion of patch-clamp data from an additional hPSC-CM line would significantly strengthen the claim that this framework can harmonize and generalize across datasets and cell sources.*

Thank you for this helpful suggestion. We agree that adding data from more hPSC-CM lines would improve the framework's generalizability. In this work, our goal was to show that the digital twin framework is data-driven and can easily be expanded to include more hPSC-CM lines, allowing for cross-line comparisons in future studies. We have clarified this and included a discussion of this limitation in the revised manuscript. We are currently seeking funding for patient-specific lines as well to allow scalability.

> *(3) The authors note that the experimental cells exhibited high variability in action potential morphology. This is an important observation that directly supports the motivation for the study and should be explicitly presented, even if only in the supplementary materials.*

Thank you for this suggestion. We agree that explicitly showing the variability in experimental action potential morphology strengthens the motivation for this study. We have now added a section in the discussion discussing this and referencing the many prior studies that focused on iPSC-CM variability, including the studies upon which our initial model (Kernik-Clancy) was based.

> *(4) In the hERG-blocker experiments, further clarification is needed regarding the biological relevance of the reported 3% incidence of early after depolarizations (EADs). Additionally, an interrupted sentence in this section makes it unclear whether the goal is to demonstrate that the digital twin can capture rare arrhythmic risk events or whether the digital twin is necessary to determine whether this level of risk is clinically meaningful.*

Thank you for this important comment. We agree that more clarification is needed on the ~3% EAD incidence and the digital-twin role. This analysis aims to show that electrophysiological variability can create a small, susceptible subpopulation under drug effects, not to set a clinical risk threshold. The observed ~3% EAD incidence reflects the emergence of such a susceptible subpopulation under hERG block. While relatively small, this fraction is important because it arises from modest, physiologically plausible variation in ionic properties and would be difficult to capture using single-cell or small-sample approaches. As described in the Discussion, this variability-driven emergence of EADs provides a quantitative measure of proarrhythmic risk at the population level. The digital-twin framework enables systematic identification and quantification of these rare events, linking cell-level variability to population-level responses. We have revised the manuscript to clarify this point.

> *(5) The manuscript states that some action potentials were excluded from the experimental dataset. A brief explanation of the exclusion criteria, along with guidance on how to distinguish high-quality from low-quality recordings, would improve transparency and reproducibility.*

Thank you for this comment. We agree that the definition of failed recordings should be clarified. We have now specified the exclusion criteria in the Methods section.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) It would be helpful if the network cartoon in Figures 2 and 3 were replaced with a simplified sketch of the actual neural network used.*

Thank you. We now have new figures 2 and 3.

> *(2) Subsection title for the Introduction has a typo.*

Thank you. We have fixed it.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) Technical quality control criteria are not specified.*
>
> *The Methods section states that "any incomplete or failed recordings were excluded," but does not define what constitutes a failed recording. The criteria could be subjective.*

Thank you for pointing this out. We agree that the definition of failed recordings should be clarified. We have now specified the exclusion criteria in the Methods section.

“Recordings were excluded if they exhibited no spontaneous firing, abnormally slow firing rates, or failed to capture a complete action potential waveform. These criteria were applied consistently across all recordings.”

> *(2) "Cell-specific" may overstate the claim.*
>
> *The term "cell-specific digital twins" (title, throughout) implies that the inferred parameters reflect the true biological state of each cell. However, parameters are derived only from curve-fitting to electrophysiological data and do not reflect other biological components (e.g., gene expression, contractility, calcium handling, metabolism, etc). Please consider rephrasing to "electrophysiology-based digital twins", "voltage clamp-matched digital twins", etc.*

Thank you for this important comment. We agree that the term “cell-specific” could be interpreted as implying a complete representation of the biological state of each cell. We have also adjusted the wording in relevant sections to avoid over-interpretation.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *(1) I would add the list of the 52 parameters in the method section/SI and not just in the reference. Additional justification of why the perturbation was set as +/- 40% for the 52 parameter or +/- 20% for the EAD population would also help.*

Thank you for this helpful comment. We have included model equations and highlighted the 52 parameters in the Supplementary Information and provided additional justification in the Methods.

> *(2) In Figure 1B, might be helpful to add the axis of the Vm instead of the dotted line indicating 0 mV to show differences in the diastolic potential.*

Thank you! We have now updated Figure 1B.

> *(3) Figure 1C-I might be more impactful to show traces from the AP shown in Figure B to reinforce the impact of a single current in the AP shape.*

We have now updated Figure 1C-I to include traces from the AP shown in Figure 1B.

https://doi.org/10.7554/eLife.110013.2.sa0