Digital Twin for the Win: Personalized Cardiac Electrophysiology

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] and Colleen E. Clancy*[1,2,5]

[1]Center for Precision Medicine and Data Science,

[2]Department of Physiology and Membrane Biology,

[3]Department of Internal Medicine, Division of Cardiovascular Medicine,

[4] Institute for Regenerative Cures,

[5]Department of Pharmacology,

University of California, Davis, California

*Corresponding author

Email: ceclancy@ucdavis.edu

CONTRIBUTIONS

P.C.Y., M.T.J., and C.C. designed the research study, conducted simulations, acquired data, and analyzed data. D.K.L. and R.S. sample preparation and differentiation of the iPSC lines. G.H.H provided experimental support. L.F.S conducted the electrophysiological recordings. P.C.Y, L.F.S. and C.E.C analyzed data, and wrote the manuscript.

Running Title: Cell-Specific iPSC-CM Digital Twins

ABSTRACT:

Individual variability drives differences in cardiovascular disease manifestation, therapeutic response, and rare phenotypes. Traditional preclinical models minimize variability, limiting their ability to capture population heterogeneity and contributing to translational gaps and adverse drug reactions. Here, we present a computational, experimental, and machine learning framework for generating digital twins of induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) that reproduce cell-specific electrophysiology from a single optimized voltage-clamp protocol. A synthetic population of iPSC-CMs was generated by varying 52 parameters across six ionic currents and used to train a neural network to infer parameters from current responses. Applied to experimental data, this approach produced digital twins that captured action potential diversity across temperatures and morphologies. Importantly, baseline variability in ionic current dynamics not only explained heterogeneity in spontaneous action potential waveforms but also predicted differential drug sensitivity, supporting digital twin application in cardiac safety pharmacology and precision medicine.

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, incomplete mechanistic understanding of disease, and a higher risk of adverse drug reactions (*4-10*). 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 (*11-17*).

Human induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) are a powerful platform for studying electrophysiology in a patient-specific context (*18-24*). However, experimental measurements of iPSC-CM action potentials (APs) or ionic currents provide only partial information about the underlying mechanisms (*25-29*). 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 data sets impossible. Existing fitting approaches are slow, require extensive specialized expertise, and are not scalable to capture individual variability (*30-37*).

Mathematical models of cardiac electrophysiology comprise a complementary framework to link cellular dynamics and underlying ionic mechanisms (*38, 39*). 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 generation of massive synthetic datasets generated from large *in silico* cell populations to generate the training data necessary for accurate and 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, cell-specific 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 models reconstructed from the model-generated parameters reproduce the AP waveform, as well as 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 the Methods section). 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 perturbation, there was extensive variability observed in the model population like what is observed 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 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 variety, we examined the total membrane current (I_{total}) and its decomposition into individual ionic components for representative cells across the population (**Figure 1C–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} amplitude with 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, cell-specific 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 synthetic population of 200,000 iPSC-CMs 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 (**Figure 2, top left panel**).

As part of the workflow (**Figure 2 top row**), a voltage-clamp protocol 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 **Figure 2 top left center panel)**.

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 data sets (**Figure 2 top center right pane**l). The model reached consistently low MSE values across both training and test datasets, indicating strong generalization. An example parameterization for the fast sodium current (I_{Na}) illustrates that the network infers gating kinetic parameters and maximal conductance (**Figure 2, top right panel**). The bottom row shows the digital twin output with inferred parameters used to simulate the AP, whole cell currents I_{Total} and each key ion channel I_{Na}, I_{CaL}, I_{K1}, I_{Ks}, I_{Kr}, and I_f as shown in **Figure 2 bottom row.**

To carry out the process to create digital twins for iPSC-CMs, we first set out to design a voltage-clamp protocol empirically, guided by our extensive historical knowledge of cardiac ion channel gating kinetics (**Figure 3A**). The protocol consisted of a complex series of voltage steps intended to activate the full complement of major ionic currents in iPSC-CMs, providing temporally distinct current signatures for parameter estimation. When applied to a computational population of iPSC-CMs with randomized parameter sets, the protocol (V_m) successfully elicited robust whole-cell current responses (I_{Total}) with clearly resolvable transient and sustained components.

Empirical Protocols and Training Data for Ion Channel Characterization

Decomposition of the simulated whole-cell currents confirmed that the protocol effectively isolated the distinct activation patterns of the major ionic species (**Figure 3B–G**). I_{CaL} was prominently engaged during depolarizing steps, I_{Kr} and I_{Ks} demonstrated rapid and delayed activation, respectively, consistent with their role in repolarization, I_{Na} produced sharp inward transients at the onset of depolarization, I_f was activated during hyperpolarizing phases, and I_{K1} contributed steady inward rectification at hyperpolarized potentials. These results confirmed that the synthetic data set subject to

the protocol resulted in rich, information-dense data well suited for deep learning-based parameter extraction.

We next evaluated the deep learning (DL) approach to extract ionic model parameters from whole cell current responses elicited by the empirically designed voltage-clamp protocol (**Figure 4A**). As described above, the voltage protocol (**A, left**) was constructed to sequentially and selectively to activate major ionic currents while capturing the kinetics of each current. When applied to computer generated synthetic iPSC-CMs, dynamic changes in membrane potential (**panel A, left, V_m, black trace**) produced complex whole-cell current responses (**panel A, right, I_{Total}, red trace**). The inset highlights the rapid transient components of the current, providing key information about the fast activation and inactivation kinetics. The full-length current traces were used as direct input for deep learning-based parameter estimation.

To investigate the influence of training data size on predictive accuracy, we trained networks with datasets containing 1,000, 10,000, and 200,000 synthetic cells (**Figure 4B**). For the smallest dataset, the test error curve exhibited an asymmetric U-shaped profile despite continued decreases in training error, suggesting overfitting and poor generalization. Increasing the training set to 10,000 cells reduced both the magnitude and variability of test error, but not within our stated error limits. The largest dataset of 200,000 cells successfully reduced prediction errors, with training and test curves converging to a median mean absolute error (MAE) of 0.023. The corresponding MAE distributions narrowed substantially with increasing dataset size, indicating improved stability and generalizability of parameter predictions.

We then assessed whether the deep learning-based estimated parameters could reproduce full-cell electrophysiological behavior. Using the network trained on the largest dataset, we simulated APs, intracellular calcium transients (Ca^{2+}), total ionic current (I_{Total}), and six individual ionic currents (I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f) for a representative test cell (**Figure 4C**). The deep learning network-generated parameters resulted in model waveforms (red) that closely overlapped the ground truth simulations (blue) for all outputs, with high fidelity in both amplitude and temporal features. This included precise reproduction of AP morphology, Ca^{2+} transient dynamics, and the activation and inactivation profiles of each ion channel.

These results demonstrate that when paired with an optimized empirically determined voltage-clamp protocol, the deep learning framework can successfully derive complete cell-specific parameter sets with high accuracy and generate digital twins that replicate underlying ionic profiles and emergent

electrophysiological behavior. At least for computational models of iPSC-CMs, the protocol provided rich, information-dense data well suited for deep learning–based parameter extraction.

However, we faced substantial problems when we attempted to implement the empirical protocol in experimental settings where we applied the protocol *in vitro* to real human iPSC-CMs (cell line iPS-6-9-9T.B). High-resolution current recordings proved difficult to obtain, primarily due to challenges in maintaining voltage control across the full voltage clamp protocol and in preserving cell viability throughout the long-duration protocol. Many cells exhibited signs of compromised seal integrity or loss of responsiveness before the protocol was completed, reducing the quality and completeness of the experimental datasets. As a result, while the empirically derived protocol was effective for synthetic data, its length and complexity limited its applicability to live cell recordings, and consequently to the construction of digital twins from those cells.

Deep Learning Approaches for Voltage Clamp Optimization and Parameter Estimation

To address these challenges, we encountered with the empirically designed voltage-clamp protocol, we concluded that a more optimized and simpler protocol was required, and that our historical human knowledge was insufficient to effectively design the best protocol. To maximize the information from experimental recordings for parameter inference, we developed a deep learning–guided optimization strategy for designing ideal voltage-clamp protocols (**Figure 5**). Starting from a broad set of candidate voltage steps, the algorithm iteratively evaluated MSE between predicted and true parameters for each segment of the voltage protocol in large synthetic iPSC-CM populations. In each optimization cycle, the most informative test potential was selected and incorporated into the evolving protocol, while less informative steps were iteratively discarded.

Each optimization cycle began with a −100 mV holding potential for 250 ms, followed by a systematic sweep through test potentials ranging from −120 mV to +50 mV in 10 mV increments. For each training cycle, a population of 200,000 computational iPSC-CMs with randomized parameter sets was simulated under this protocol, and the resulting whole-cell current responses were used to train the deep learning model to predict all 52 ionic parameters.

The performance of each test membrane potential was evaluated by computing the MSE between predicted and true parameters. The test membrane potential producing the lowest MSE was then selected as the optimal step for that cycle and applied for 250 ms before resuming the sequence. This process formed an iterative loop of evaluation, MSE-based selection, and protocol updating, repeated

every 7000 ms. At 6000 ms, a transient step to −120 mV for 250 ms was included to probe hyperpolarization-activated channels before continuing to the next cycle.

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.

When applied to synthetic iPSC-CMs held back test data, the optimized voltage clamp protocol (**Figure 6A, left)** produced digital whole-cell currents (**Figure 6A, 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 improvement for trained networks with datasets containing increasing 1,000, 10,000, and 200,000 synthetic cells (**Figure 6B**). For the smallest dataset (1,000 cells), the test error plateaued at a 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 (**6B, left column**). 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 (**6B, middle column**). 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 0.020 (**6B, right column**). Again, the new protocol resulting in corresponding MAE distributions that narrowed substantially with increasing dataset size, indicating improved stability and generalizability of parameter predictions.

The automated deep learning network-generated parameters led to improved model waveforms (red) compared to the earlier empirical protocol, and duplicated ground truth simulated data (blue) for all outputs, with high fidelity in both amplitude and temporal features (**Figure 6C**). 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 cell-specific ionic properties is critical.

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 7**).

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 live cells in response to the protocol described in Figure 5 are shown in **Figure 7A and B** from two live iPSC-CMs served as the sole input to the trained deep learning network, which successfully inferred all 52 biophysical parameters governing all 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.

When simulated, the digital twins reproduced the experimentally recorded APs with excellent agreement, capturing fine-scale features of depolarization rate, plateau amplitude and duration, repolarization slope, and beat-to-beat variability (**Figure 7C–D**), confirming that the pipeline reconstructs the full mechanistic basis of cell-specific electrophysiology from a single experimental protocol.

To evaluate cell-specific susceptibility to drug-induced proarrhythmia, we simulated the effects of the hERG channel blocker E-4031 in two digital twin models of iPSC-CMs (**Figure 8**). In control conditions, both **Cell 1** (corresponding to Figure 7A) and **Cell 2** (corresponding to Figure 7B) exhibited stable action potentials at physiological temperature (37 °C). Upon exposure to E-4031, **Cell 1** developed early afterdepolarizations (EADs) at 50 nM, whereas **Cell 2** only exhibited EADs at the higher concentration

of 80 nM. These findings illustrate inter-cellular variability in proarrhythmic drug response, underscoring the utility of digital twin models for predicting patient-specific risk.

Building on the drug response simulations in Figure 8, we next examined how variability in network-predicted parameter sets contributes to baseline electrophysiological differences across cells (**Figure 9**). **Cell 1** (corresponding to Figure 7A) and **Cell 2** (corresponding to Figure 7B) exhibited distinct action potential morphologies that arose from divergent ionic current dynamics. In panels A and B, both **Cell 1** and **Cell 2** generated spontaneous action potentials, but differences in ionic current dynamics produced distinct waveforms. **Cell 1** exhibited a higher upstroke velocity, consistent with stronger I_{Na} contributions, whereas **Cell 2** showed reduced I_{Na} amplitude and slower depolarization. In both cells, I_{caL}, I_{Kr}, I_{K1}, and I_f shaped repolarization and diastolic depolarization**.** Panel C summarizes quantitative metrics, confirming that **Cell 1** had a markedly greater maximum upstroke velocity (145.6 vs. 21.7 mV/ms), while peak potentials and spontaneous beating rates were comparable between the two cells. These simulations demonstrate how heterogeneity in predicted parameters manifests as diverse action potential waveforms across iPSC-CMs.

This result establishes that the combination of optimized experimental design, large-scale synthetic training, and deep learning–based parameter inference enables the development of complete, cell-specific digital twins directly from experimental data. The scale of the synthetic training set, the precision of the parameter recovery, and the reproducibility across temperature and waveform diversity position this approach as a generalizable, physiologically grounded platform for personalized cardiomyocyte modeling with broad translational potential.

DISCUSSION

This study demonstrates, for the first time, the development of cell-specific digital twins of human iPSC-CMs directly from a single experimental recording. 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 reproduce its 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 conditions(*40-45*). 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(*46-49*). The framework presented here is well-suited to account for variability, providing a foundation for future studies that link experimental heterogeneity to underlying ionic mechanisms and for developing more robust, generalizable digital twin models.

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 200,000 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 initially designed a voltage-clamp protocol empirically, drawing on our extensive historical knowledge of cardiac ion channel gating kinetics (**Figure 3**). The protocol consisted of a complex series of voltage steps intended to activate the full complement of major ionic currents in iPSC-CMs, providing temporally distinct current signatures for parameter estimation. When applied to a computational population of iPSC-CMs with randomized parameter sets, the protocol successfully elicited robust whole-cell current responses (I_{Total}) with clearly resolvable transient and sustained components. The time course of these simulated currents served as the direct input for the deep learning model, enabling accurate recovery of all 52 underlying ionic parameters from both training and test datasets.

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 individual single-cell level.

However, when implemented experimentally in human iPSC-CMs, the empirically derived protocol was not successful. High-resolution current recordings proved difficult to obtain, primarily due to challenges in maintaining voltage control across the full sequence and in preserving cell viability throughout the long-duration protocol. Many cells exhibited signs of compromised seal integrity or loss of responsiveness before the protocol was completed, reducing the quality and completeness of the experimental datasets. As a result, while the empirically derived protocol was effective for synthetic data, its length and complexity limited its applicability to live cell recordings. Our experience with the empirically designed voltage-clamp protocol made it clear that relying solely on what we, as humans, know from past studies is not enough to create the best protocol.

To overcome these limitations, we transitioned to a data-driven protocol design process guided by deep learning as shown in **Figure 5**. 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 are 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 refine the protocol in real time. 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 deep learning–guided protocol optimization framework described here lays the groundwork for the capstone result presented in **Figure 7**, 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.

The simulations presented in **Figures 8 and 9** highlight two complementary aspects of variability in iPSC-CM electrophysiology: drug sensitivity and intrinsic baseline heterogeneity. In response to the hERG channel blocker E-4031 (**Figure 8**), the two digital twin models displayed markedly different thresholds for developing early afterdepolarizations (EADs). Whereas **Cell 1** exhibited EADs at 50 nM, **Cell 2** only developed EADs at 80 nM, underscoring how even subtle differences in underlying electrophysiological properties can translate into distinct proarrhythmic outcomes. This finding reflects the clinical observation that susceptibility to drug-induced arrhythmias is highly patient-specific, and it emphasizes the potential of digital twin modeling to capture and predict such individualized responses.

Together, these results demonstrate that baseline variability in ionic current dynamics not only contributes to heterogeneity in spontaneous action potential waveforms but also shapes differential responses to pharmacological interventions. By linking parameter predictions from a network model to drug sensitivity and arrhythmia susceptibility, our approach supports the use of iPSC-derived digital twins as a framework for precision medicine applications in cardiac safety pharmacology and patient-specific risk stratification.

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 examples spanning the full range of plausible ionic parameter combinations and electrophysiological phenotypes. Generating this scale and diversity experimentally would be prohibitively time consuming, costly, and technically unfeasible, given the complexity of acquiring high-fidelity voltage-clamp data and the variability inherent in biological preparations. 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, enhancing the network's ability to generalize to diverse experimental cells. By using synthetic data as the foundation for learning, the

method gains both the scale and breadth required for robust, high-fidelity digital twin modeling that would be unattainable with experimental data alone.

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 developed and validated *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 other 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.

In the near term, this approach can enable high throughput cell specific phenotyping for disease modeling, allowing investigators to map the electrophysiology of patient derived cardiomyocytes to individualized ionic mechanisms using a single experiment. Such mechanistic insight can guide precision medicine strategies including the design of pharmacological interventions tailored to the unique ionic profiles of individuals. The framework maintains accuracy across changes in temperature and to accommodate wide variation in AP morphology makes it suitable for multicenter studies in which experimental conditions differ, thereby enhancing reproducibility and enabling harmonization of data. In preclinical safety assessment, temperature independent digital twins could be used to evaluate the effects of new compounds 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

HiPSC-derived Ventricular Cardiomyocyte Differentiation and Culture

The hiPSC lines iPS-6-9-9T.B were purchased from WiCell and cultured in StemMACS iPS-Brew XF medium (Miltenyi Biotec) on hESC-qualified Matrigel (Corning) at 37ºC with 5% CO_2 and passaged at ~70% confluency. Ventricular cardiomyocytes (VCMs) were differentiated by treating confluent hiPSCs with 6 µM CHIR99021 (Tocris) from day 0 to 1 (D0–D1) and 5 µM IWR-1 (Tocris) from D2–D4 in RPMI 1640 supplemented with B27 without insulin (Thermo Fisher). From D7 onward, hiPSC-derived VCMs were cultured in cardiomyocyte (CM) maintenance medium consisting of RPMI 1640 with B27 supplement with insulin. The hiPSC-VCMs were replated on ~D14 and allowed to proliferate in CM maintenance medium with 2 µM CHIR99021 for ~7 days until they were replated onto Matrigel-coated glass coverslips for patch-clamp experiments.

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.

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 5) 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 experimental acquisition rates. 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 detailed in (*50*)) 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 slowed all channel kinetics in the synthetic training population using a Q_{10} factor of 3. Two synthetic populations of model cells were generated. The first consisted of 200,000 cells simulated at 310 Kelvin (physiological temperature), and the second included 1,100,000 cells simulated at 298 K (room temperature). In both populations, uniformly distributed random perturbations of ±40% were applied to each model parameter relative to baseline values, ensuring physiologically plausible variability. The training dataset was synthetically generated; no missing data occurred.

Fully connected layers

Fully connected deep neural networks with twenty-six double precision hidden layers (1024 nodes; tanh activation) were implemented in TensorFlow. Every neuron in a layer was connected to neurons in the next layer (*51*). The fully connected layers calculated a weighted sum of I_{Total} and added a bias term to the outputs. These data were then passed to a non-linear activation function (*f*) to define the output for each neuron (Eqs. 1 and 2) (*52*).

$$y_j^k = f\left(Z_j^k\right) \quad (1)$$

$$Z_j^k = W_{i,j}^k * a_j^{k-1} + b^k \quad (2)$$

Where $k \in \{1, \dots, n\}$, and (*i* , *j)* represent the number of hidden layers and neurons in each pair of subsequent hidden layers. The optimized values for these parameters were found via hyperparameter tuning where a^{k} is each neuron output where $a^0 \in \{IC_1, \dots, IC_m\}$ is the input to the fully connected layers, and a^{n+1} is the network output: $\widehat{p}$ is the outputs for parameters. We first assigned random values to all network parameters θ_t; each neuron weight $\left(W_{i,j}\right)$, bias term (b), which is a constant added to calculate the neurons output and other network hyperparameters (the number of hidden

layers, the number of neurons for each hidden layer and activation functions for each hidden layer) to start the optimization process for finding the best network infrastructure.

Network performance

Networks were trained using mean squared error (MSE) loss and the Adam optimizer (1×10^{-4}). MSE-loss of function evaluation metric (Eq. 3) to predicted model parameters (*53, 54*).

$$MSE = \frac{1}{m} \sum_{i=1}^{n} \|p_i - \hat{p}_i\|^2 \qquad (3)$$

where *m* is the total number of whole-cell membrane current plus calcium transient (IC_m) and p_i and $\hat{p}_i$ are the corresponding parameters and its predicted parameters (the network output), with respect to the sample cells. We used a hyperbolic tangent (tanh) as activation function in Eq. 1 to calculate the output for each hidden layer neuron at each training iteration.

Training datasets of 1,000, 10,000, 200,000, 1,100,000 cells were independently sampled from the full population to evaluate the impact of dataset size on predictive performance. The dataset size (1.1M cells) was selected to ensure broad exploration of parameter space. Subsets of 1,000, 10,000, and 200,000 cells were sampled to evaluate learning curve effects and to demonstrate the impact of dataset size on predictive accuracy. Because the training dataset was synthetically generated, no missing data occurred.

We split the data into training, validation, and test sets using a 70:10:20 ratio, and implemented the network architecture using Keras (*55*). The dataset was also divided into 1% sample batches, requiring 70 iterations to complete one epoch. As we monitor 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 the evaluation metrics using validation data during each iteration of training and compare to values from training data. At the point when model performance on the training data began to degrade compared 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 traces of ionic currents and action potentials (APs) closely mirror the physiological behavior of individual cardiac cells.

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 follow:

$$G_{Kr} = G_{Kr,max} * \left(\frac{1}{1 + \ ([E - 4301]/IC_{50})^{n}}\right)$$

where $G_{Kr,max}$ is the nominal conductance value obtained from each ventricular myocyte model, and the IC_{50} (7 nM) (*50, 56*) 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.

DATA AVAILABILITY

Experimental dataset generated for this study is available at https://github.com/ClancyLabUCD/Digital-Twin-for-the-Win-Personalized-Cardiac-Electrophysiology).

CODE AVAILABILITY

Due to the large size of the synthetic datasets (~1.1M cells), the raw data are not hosted online. However, a representative a subset 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.

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.

DISCLOSURES

D.K. Lieu serves as a scientific consultant for Novoheart, Ltd.

REFERENCE

1. R. Serrano *et al.*, A deep learning platform to assess drug proarrhythmia risk. *Cell Stem Cell* **30**, 86-95 e84 (2023).
2. G. Gintant, P. T. Sager, N. Stockbridge, Evolution of strategies to improve preclinical cardiac safety testing. *Nat Rev Drug Discov* **15**, 457-471 (2016).
3. H. Laverty *et al.*, How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? *Br J Pharmacol* **163**, 675-693 (2011).
4. J. Seok *et al.*, Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A* **110**, 3507-3512 (2013).
5. A. Akhtar, The flaws and human harms of animal experimentation. *Camb Q Healthc Ethics* **24**, 407-419 (2015).
6. M. B. Bracken, Why animal studies are often poor predictors of human reactions to exposure. *J R Soc Med* **102**, 120-122 (2009).
7. K. Fabre *et al.*, Introduction to a manuscript series on the characterization and use of microphysiological systems (MPS) in pharmaceutical safety and ADME applications. *Lab Chip* **20**, 1049-1057 (2020).
8. R. Franco, A. Cedazo-Minguez, Successful therapies for Alzheimer's disease: why so many in animal models and none in humans? *Front Pharmacol* **5**, 146 (2014).
9. E. K. Heist, J. N. Ruskin, Drug-induced arrhythmia. *Circulation* **122**, 1426-1435 (2010).
10. P. K. Sorger, S. R. B. Allerheiligen, Quantitative and Systems Pharmacology in the Post-genomic Era: New Approaches to Discovering Drugs and Understanding Therapeutic Mechanisms. *Tech. Rep.*, (2011).
11. M. Fink *et al.*, Cardiac cell modelling: observations from the heart of the cardiac physiome project. *Prog Biophys Mol Biol* **104**, 2-21 (2011).
12. P. C. Yang *et al.*, Toward Digital Twin Technology for Precision Pharmacology. *JACC Clin Electrophysiol* **10**, 359-364 (2024).
13. E. Passini *et al.*, Human In Silico Drug Trials Demonstrate Higher Accuracy than Animal Models in Predicting Clinical Pro-Arrhythmic Cardiotoxicity. *Front Physiol* **8**, 668 (2017).
14. D. C. Kernik, P. C. Yang, J. Kurokawa, J. C. Wu, C. E. Clancy, A computational model of induced pluripotent stem-cell derived cardiomyocytes for high throughput risk stratification of KCNQ1 genetic variants. *PLoS Comput Biol* **16**, e1008109 (2020).

15. A. Tveito, K. H. Jaeger, M. M. Maleckar, W. R. Giles, S. Wall, Computational translation of drug effects from animal experiments to human ventricular myocytes. *Sci Rep* **10**, 10537 (2020).
16. A. Tveito *et al.*, Inversion and computational maturation of drug response using human stem cell derived cardiomyocytes in microphysiological systems. *Sci Rep* **8**, 17626 (2018).
17. S. Langthaler *et al.*, A549 in-silico 1.0: A first computational model to simulate cell cycle dependent ion current modulation in the human lung adenocarcinoma. *PLoS Comput Biol* **17**, e1009091 (2021).
18. D. K. Lieu *et al.*, Mechanism-based facilitated maturation of human pluripotent stem cell-derived cardiomyocytes. *Circ Arrhythm Electrophysiol* **6**, 191-201 (2013).
19. A. De La Mata *et al.*, BIN1 Induces the Formation of T-Tubules and Adult-Like Ca(2+) Release Units in Developing Cardiomyocytes. *Stem Cells* **37**, 54-64 (2019).
20. D. Thomas, S. Shenoy, N. Sayed, Building Multi-Dimensional Induced Pluripotent Stem Cells-Based Model Platforms to Assess Cardiotoxicity in Cancer Therapies. *Front Pharmacol* **12**, 607364 (2021).
21. T. de Korte *et al.*, 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 (2020).
22. A. P. Hnatiuk, F. Briganti, D. W. Staudt, M. Mercola, Human iPSC modeling of heart disease for drug development. *Cell Chem Biol* **28**, 271-282 (2021).
23. N. Sayed, C. Liu, J. C. Wu, Translation of Human-Induced Pluripotent Stem Cells: From Clinical Trial in a Dish to Precision Medicine. *J Am Coll Cardiol* **67**, 2161-2176 (2016).
24. P. W. Burridge *et al.*, Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity. *Nat Med* **22**, 547-556 (2016).
25. A. Bizy, M. Klos, Optimizing the Use of iPSC-CMs for Cardiac Regeneration in Animal Models. *Animals (Basel)* **10**, (2020).
26. P. Wu *et al.*, Maturation strategies and limitations of induced pluripotent stem cell-derived cardiomyocytes. *Biosci Rep* **41**, (2021).
27. S. Vuckovic *et al.*, Characterization of cardiac metabolism in iPSC-derived cardiomyocytes: lessons from maturation and disease modeling. *Stem Cell Res Ther* **13**, 332 (2022).
28. H. S. Hwang *et al.*, Comparable calcium handling of human iPSC-derived cardiomyocytes generated by multiple laboratories. *J Mol Cell Cardiol* **85**, 79-88 (2015).
29. C. C. Veerman *et al.*, Immaturity of human stem-cell-derived cardiomyocytes in culture: fatal flaw or soluble problem? *Stem Cells Dev* **24**, 1035-1052 (2015).
30. M. Fink, D. Noble, Markov models for ion channels: versatility versus identifiability and speed. *Philos Trans A Math Phys Eng Sci* **367**, 2161-2179 (2009).
31. W. Groenendaal *et al.*, Cell-specific cardiac electrophysiology models. *PLoS Comput Biol* **11**, e1004242 (2015).
32. A. X. Sarkar, E. A. Sobie, Regression analysis for constraining free parameters in electrophysiological models of cardiac cells. *PLoS Comput Biol* **6**, e1000914 (2010).
33. J. D. Moreno, T. J. Lewis, C. E. Clancy, Parameterization for In-Silico Modeling of Ion Channel Interactions with Drugs. *PLoS One* **11**, e0150761 (2016).
34. T. Guo, A. Al Abed, N. H. Lovell, S. Dokos, Optimisation of a generic ionic model of cardiac myocyte electrical activity. *Comput Math Methods Med* **2013**, 706195 (2013).
35. J. Nelder, R. Mead, A simplex method for function minimization. *Computer Journal* **7**, 308-313 (1965).
36. S. Dokos, N. H. Lovell, Parameter estimation in cardiac ionic models. *Prog Biophys Mol Biol* **85**, 407-431 (2004).
37. B. B. Hui, S. Dokos, N. H. Lovell, 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 (2007).
38. C. L. Lei *et al.*, 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**, 20190348 (2020).

39. A. P. Clark, S. Wei, K. Fullerton, T. Krogh-Madsen, D. J. Christini, Single-cell ionic current phenotyping explains stem cell-derived cardiomyocyte action potential morphology. *Am J Physiol Heart Circ Physiol* **326**, H1146-H1154 (2024).
40. G. E. Kirsch *et al.*, Variability in the measurement of hERG potassium channel inhibition: effects of temperature and stimulus pattern. *J Pharmacol Toxicol Methods* **50**, 93-101 (2004).
41. J. A. Yao *et al.*, Estimation of potency of HERG channel blockers: impact of voltage protocol and temperature. *J Pharmacol Toxicol Methods* **52**, 146-153 (2005).
42. J. T. Milnes, H. J. Witchel, J. L. Leaney, D. J. Leishman, J. C. Hancox, Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 degrees C: Cisapride versus dofetilide. *J Pharmacol Toxicol Methods* **61**, 178-191 (2010).
43. K. H. Narsinh *et al.*, Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J Clin Invest* **121**, 1217-1221 (2011).
44. M. X. Doss *et al.*, Maximum diastolic potential of human induced pluripotent stem cell-derived cardiomyocytes depends critically on I(Kr). *PLoS One* **7**, e40288 (2012).
45. D. T. Du, N. Hellen, C. Kane, C. M. Terracciano, 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 (2015).
46. O. J. Britton *et al.*, Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proc Natl Acad Sci U S A* **110**, E2098-2105 (2013).
47. J. Q. X. Gong, E. A. Sobie, Population-based mechanistic modeling allows for quantitative predictions of drug responses across cell types. *NPJ Syst Biol Appl* **4**, 11 (2018).
48. H. Ni, S. Morotti, E. Grandi, A Heart for Diversity: Simulating Variability in Cardiac Arrhythmia Research. *Front Physiol* **9**, 958 (2018).
49. A. Muszkiewicz *et al.*, 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 (2016).
50. D. C. Kernik *et al.*, A computational model of induced pluripotent stem-cell derived cardiomyocytes incorporating experimental variability from multiple data sources. *J Physiol* **597**, 4533-4564 (2019).
51. A. Krogh, What are artificial neural networks? *Nature biotechnology* **26**, 195-197 (2008).
52. O. Carugo, F. Eisenhaber, Carugo, *Data mining techniques for the life sciences*. (Springer, 2010), vol. 609.
53. I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*. (MIT press, 2016).
54. K. P. Murphy, *Machine learning: a probabilistic perspective*. (MIT press, 2012).
55. K. B. Prakash, R. Kannan, S. A. Alexander, G. R. Kanagachidambaresan, *Advanced Deep Learning for Engineers and Scientists: A Practical Approach*. (Springer, Switzerland, 2021).
56. K. Harris *et al.*, Comparison of electrophysiological data from human-induced pluripotent stem cell-derived cardiomyocytes to functional preclinical safety assays. *Toxicol Sci* **134**, 412-426 (2013).

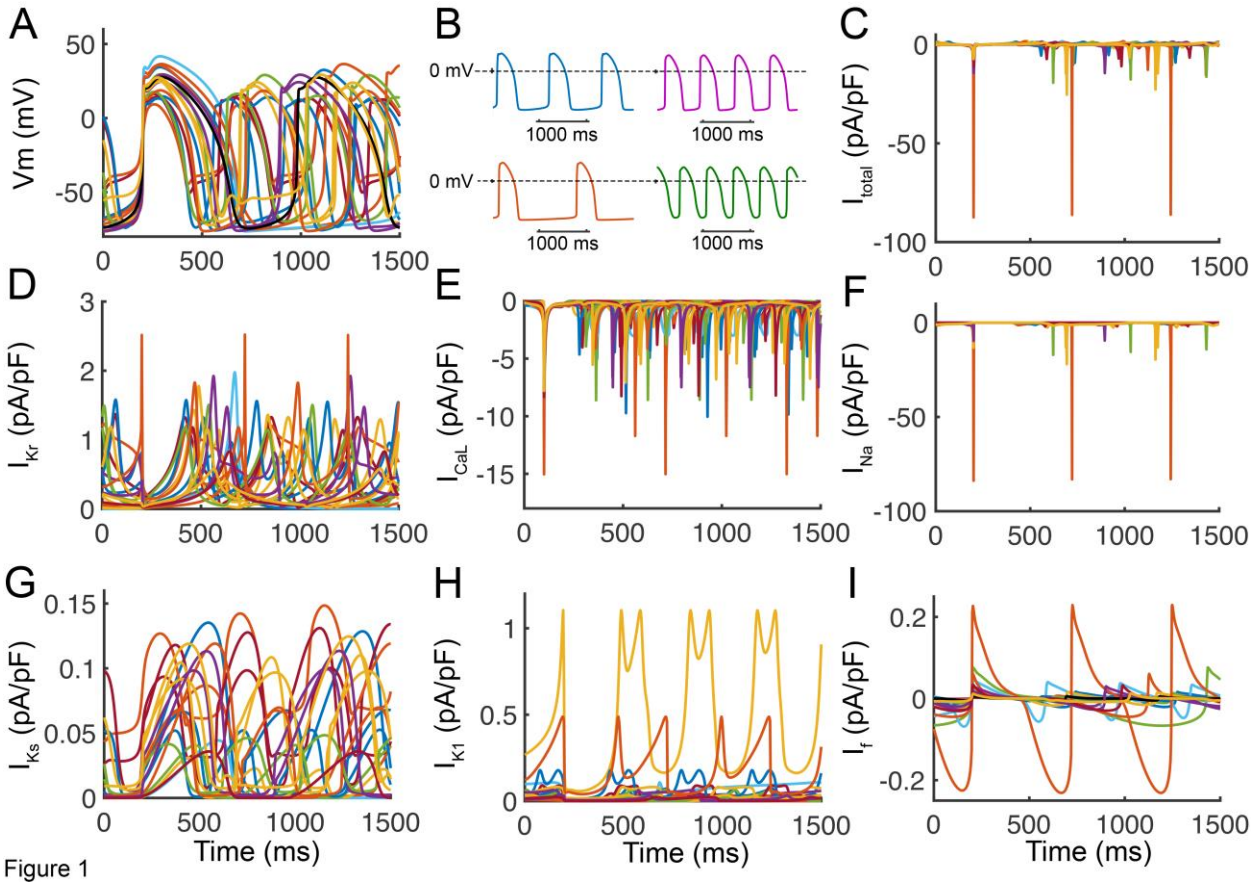


Figure 1: The variability of APs in simulated iPSC-derived cardiomyocyte populations. (A) To illustrate population variability, 20 APs were shown, each resulting from ±40% random variation applied to 52 parameters governing six key ionic currents (I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f,) in the baseline Kernik iPSC-CM model[50], 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. **(C)** The corresponding total ionic current and its decomposition into individual current components are visualized in panels **(**I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f, **D–I, respectively)**.

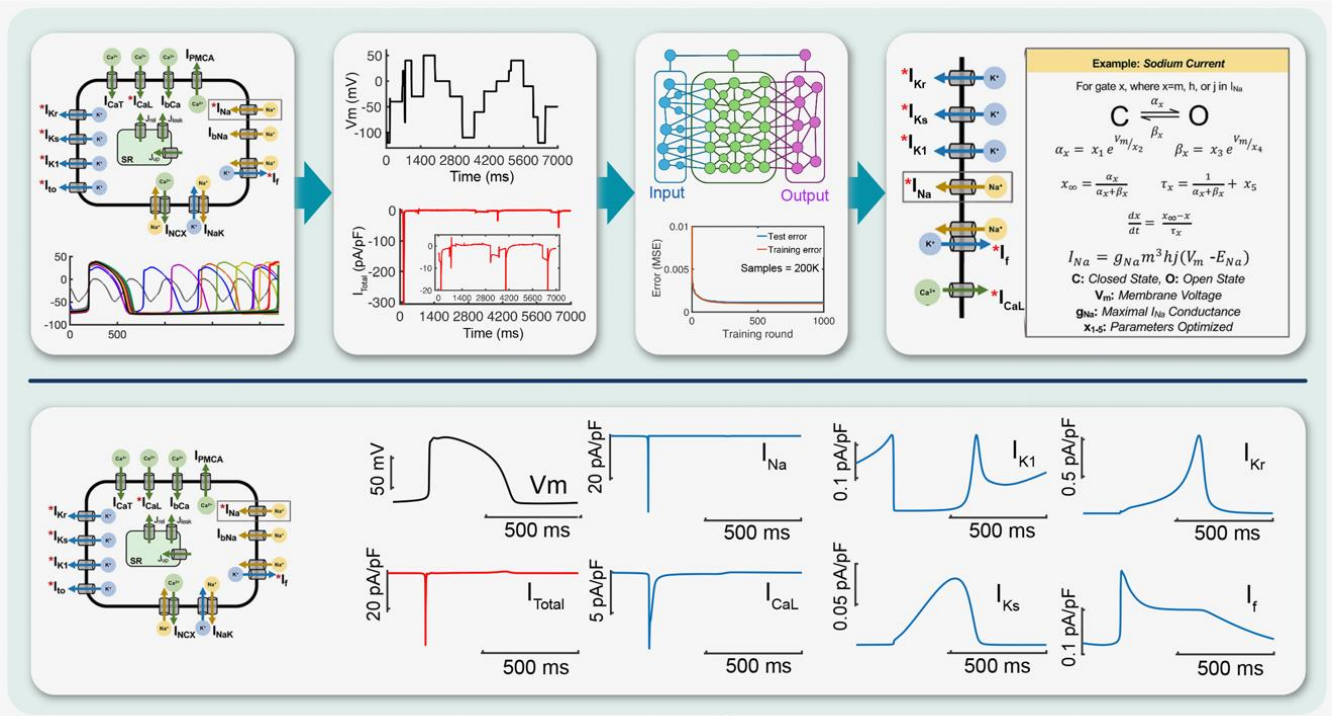


Figure 2

Figure 2. Digital twins of human iPSC-CMs from one simple voltage-clamp recording. Top row (pipeline overview): A large population of synthetic iPSC-CMs (top left panel) 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 right). A computationally optimized voltage-clamp protocol (black trace, top left center panel) is applied to generate distinct whole cell current (I_{Total}, red trace), enabling cell-wide excitation of key ion channels. The simulated whole cell currents I_{Total} from 200,000 synthetic cells serve as inputs to a fully connected neural network, trained to map raw current responses to the underlying 52 model parameters with high accuracy, as demonstrated by the low MSE across training and test sets (top center right panel). 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 (top right panel). In the **bottom row**, the resultant digital twin output is shown with the inferred parameters used to simulate the AP, whole cell currents I_{Total} and each key ion channel I_{Na}, I_{CaL}, I_{K1}, I_{Ks}, I_{Kr}, and I_f.

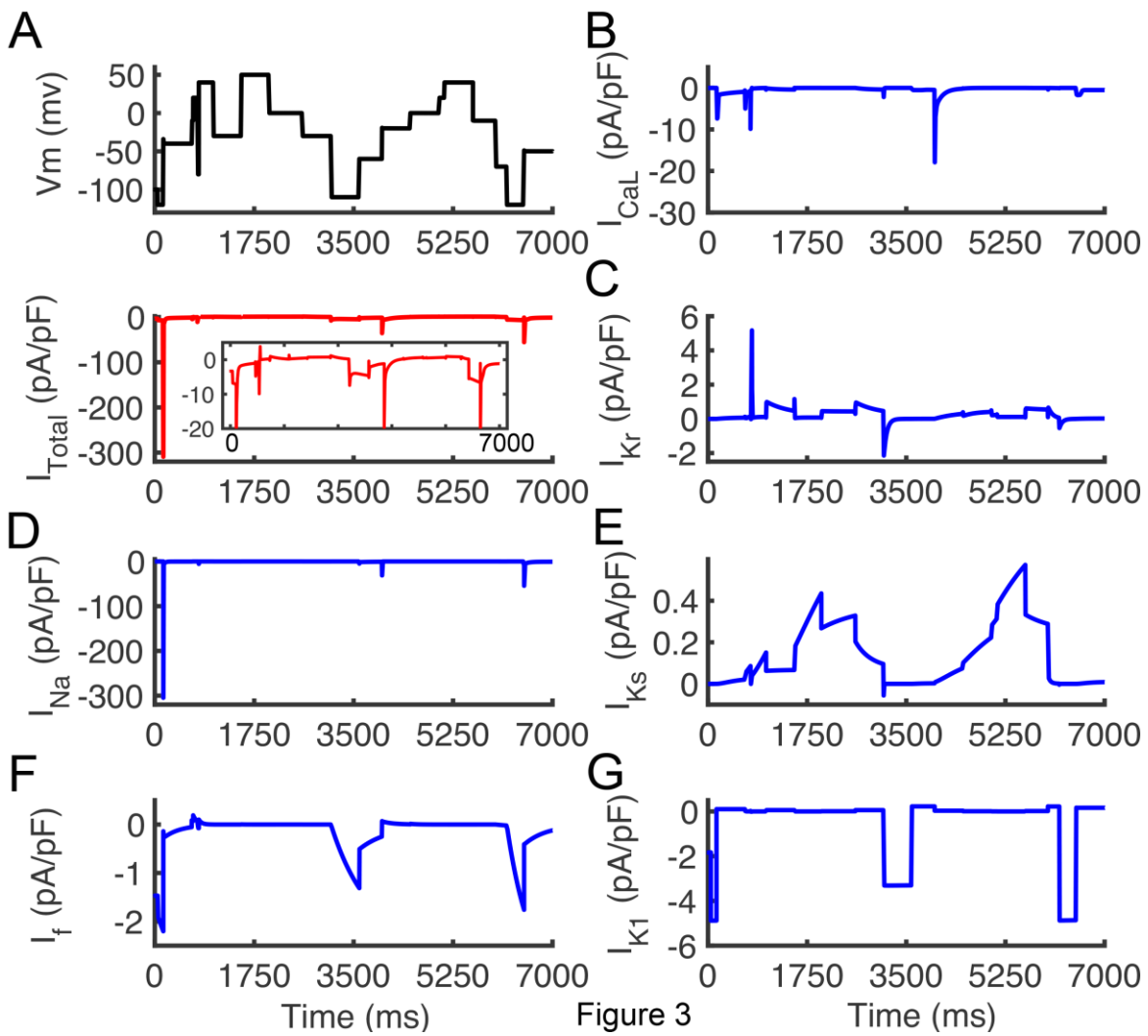


Figure 3: Simulated training dataset for parameter extraction using an empirical voltage clamp protocol. (A) The empirical membrane voltage protocol (Vm, top) applied to *in silico* iPSC-CMs. The bottom panel shows the resulting whole-cell current response (red), with an inset providing an enlarged view of the current trace for better visualization of transient events. The time course of the whole-cell current was used as input for a deep learning model to extract key ion channel parameters. (B–G) Simulated individual ionic currents generated during the same protocol: (B) L-type calcium current I_{CaL}, (C) rapid delayed rectifier potassium current I_{Kr}, (D) fast sodium current I_{Na}, (E) slow delayed rectifier potassium current I_{Ks}, (F) hyperpolarization-activated funny current I_f, and (G) inward rectifier potassium current I_{K1}. Each trace demonstrates the distinct temporal activation patterns and amplitudes of the respective ionic species under the applied voltage protocol.

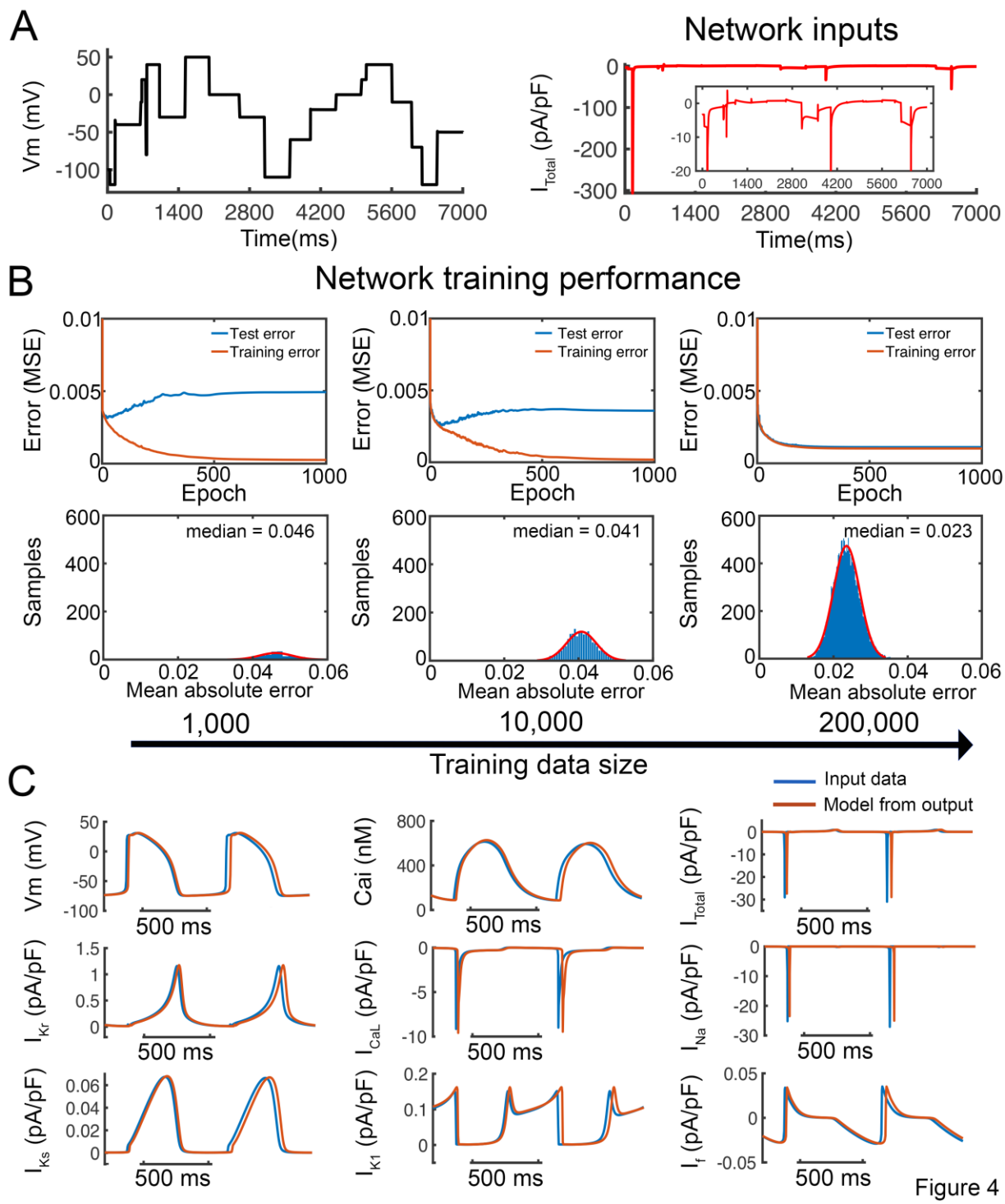


Figure 4: An empirical protocol for estimating ion channel parameters. (A) A voltage-clamp protocol is applied to induce characteristic current responses (I_{total} red trace, right) in response to dynamic changes in membrane potential (V_m, black trace), effectively stimulating key ion channels across the cell. The inset highlights the transient

components of the current response. (B) Prediction error for each model parameter was evaluated using training datasets of sizes 1,000, 10,000, and 200,000 (left to right). For smaller datasets, the test error exhibited an asymmetric U-shaped curve, while the training error consistently decreased over iterative cycle (Epoch). Mean absolute error (MAE) distributions for each dataset size are shown in the bottom panels, with median MAE values indicated. Increasing training data size reduced variability and improved agreement between training and test errors, indicating enhanced model generalization. (C) Simulated APs, intracellular calcium transients (Ca^{2+}), whole-cell current, and six individual ionic currents; I_{Kr}, I_{CaL}, I_{Na}, I_{Ks}, I_{K1}, and I_f, generated using single-cell model parameters predicted by the deep learning (DL) network trained on 200,000 data points (lowest MAE case from panel B, right). Blue traces represent ground truth input data, and red traces represent deep learning model predictions, shown over a 1600 ms time window.

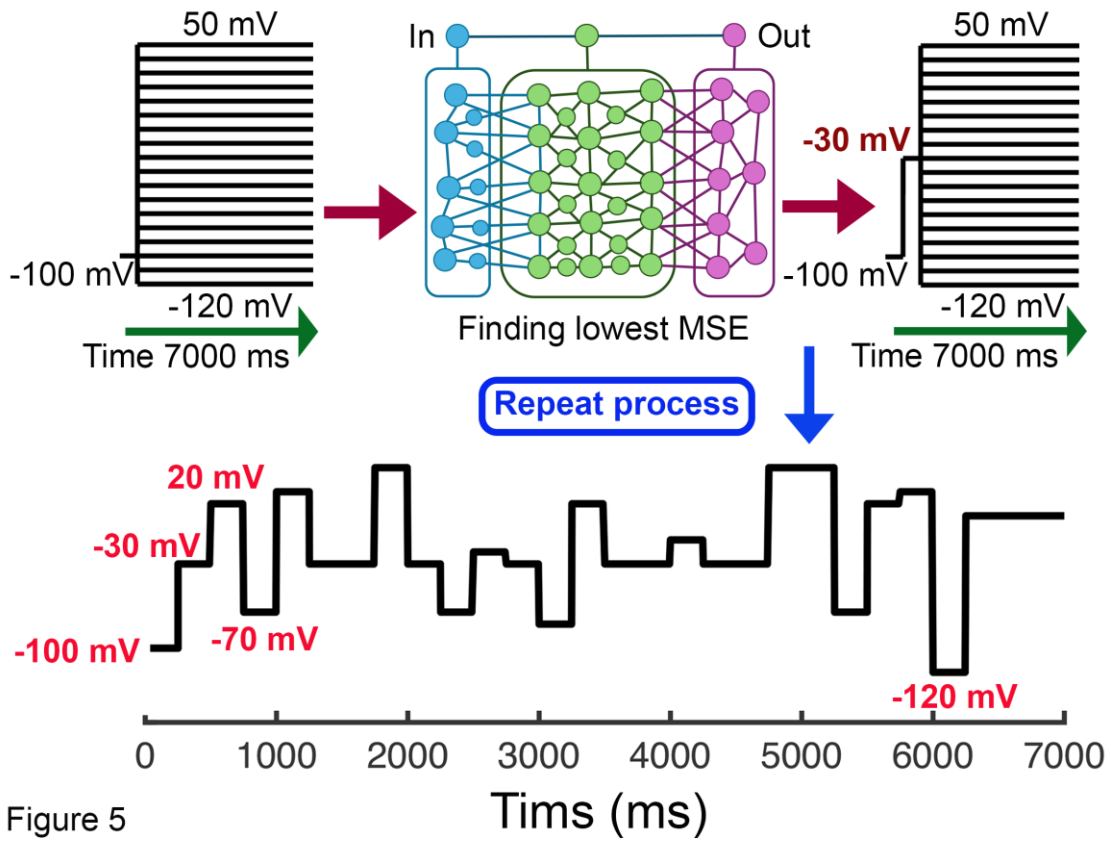


Figure 5: Deep learning guided optimization of an ideal voltage clamp protocol. Each iteration began with a −100 mV holding potential for 250 ms, followed by sequential testing potentials from −120 mV to +50 mV in 10 mV increments. A total of 200,000 synthetic samples were generated per training cycle. 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 waveform shown in the lower panel.

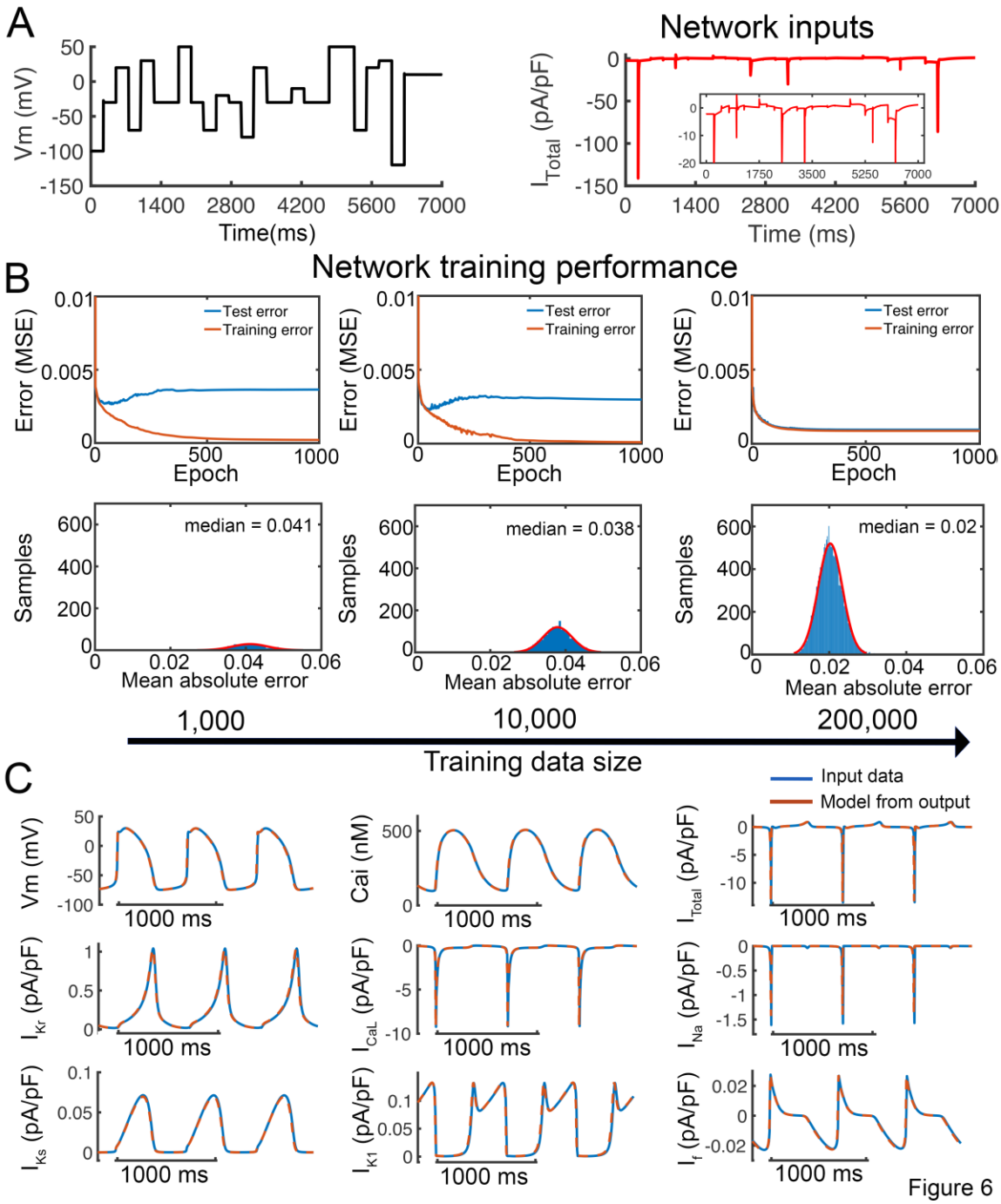


Figure 6. Deep learning for ion channel parameter estimation. (A) A deep learning derived optimized voltage-clamp protocol (left panel) was designed to activate a broad set of ionic currents in iPSC-CMs by applying dynamic membrane potential steps (black trace), producing distinctive whole-cell current responses (I_{Total}, red trace). The inset highlights fine-scale current kinetics captured during the protocol. These V_m-I_{Total} pairs

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). For the smallest dataset, the test error exhibited an asymmetric U-shaped curve across training epochs, indicating limited generalizability. As 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 with larger training datasets. **(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) 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 accurate parameter recovery and faithful reproduction of cell-specific electrophysiological behavior.

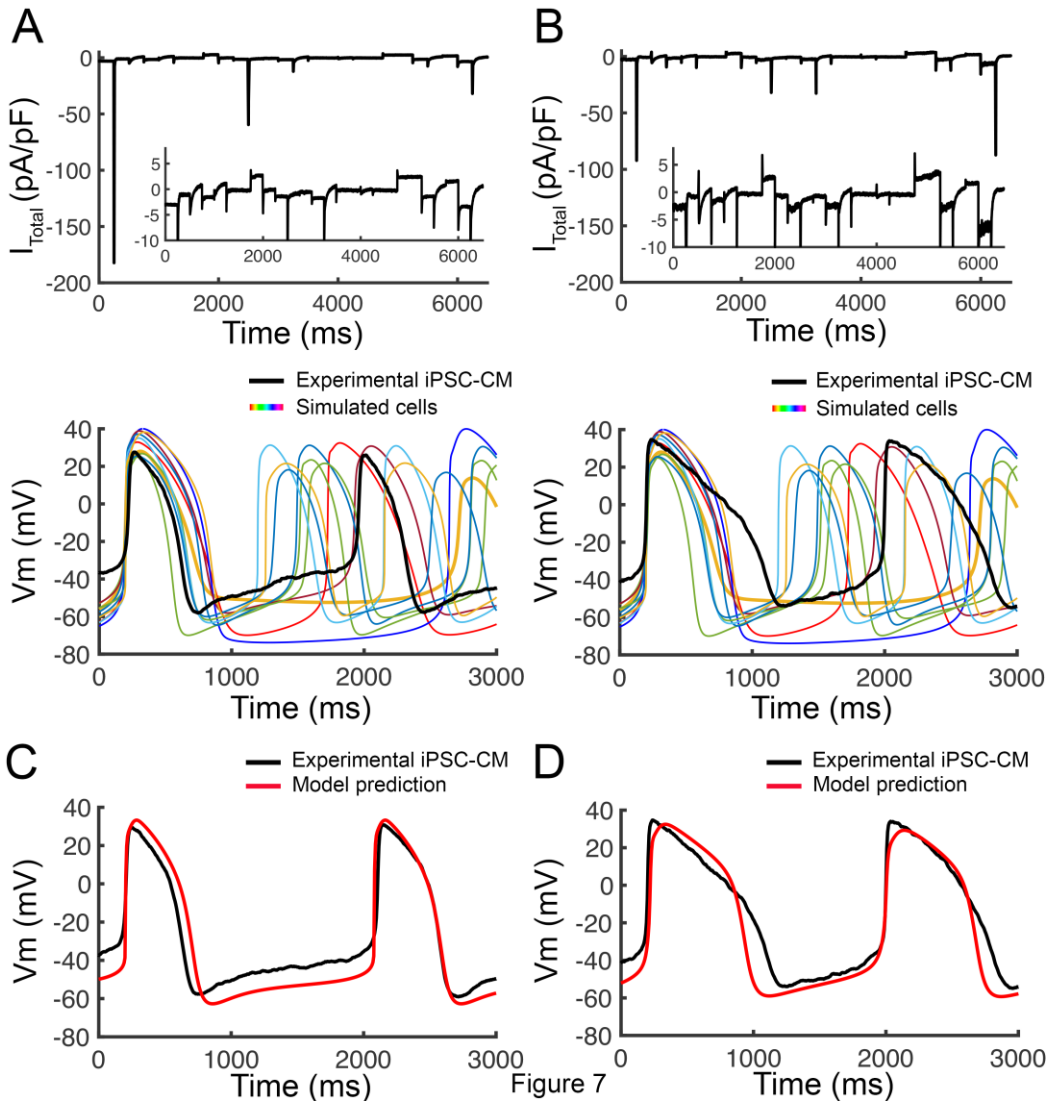


Figure 7. Digital twin generation and predictive modeling of iPSC-CM APs. (A, B, top) Whole-cell currents recorded from live cells in response to the protocol described in Figure 5. **(A, B, Middle)** Ten simulated APs from a synthetic population of 1,100,000 induced pluripotent stem cell-derived cardiomyocytes (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 from hiPSC cell line iPS-6-9-9T.B. Insets show the experimental whole-cell current responses to a deep learning optimized voltage-clamp protocol (as described in Figure 5), used for

parameter inference. **(C, D, bottom)** 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.

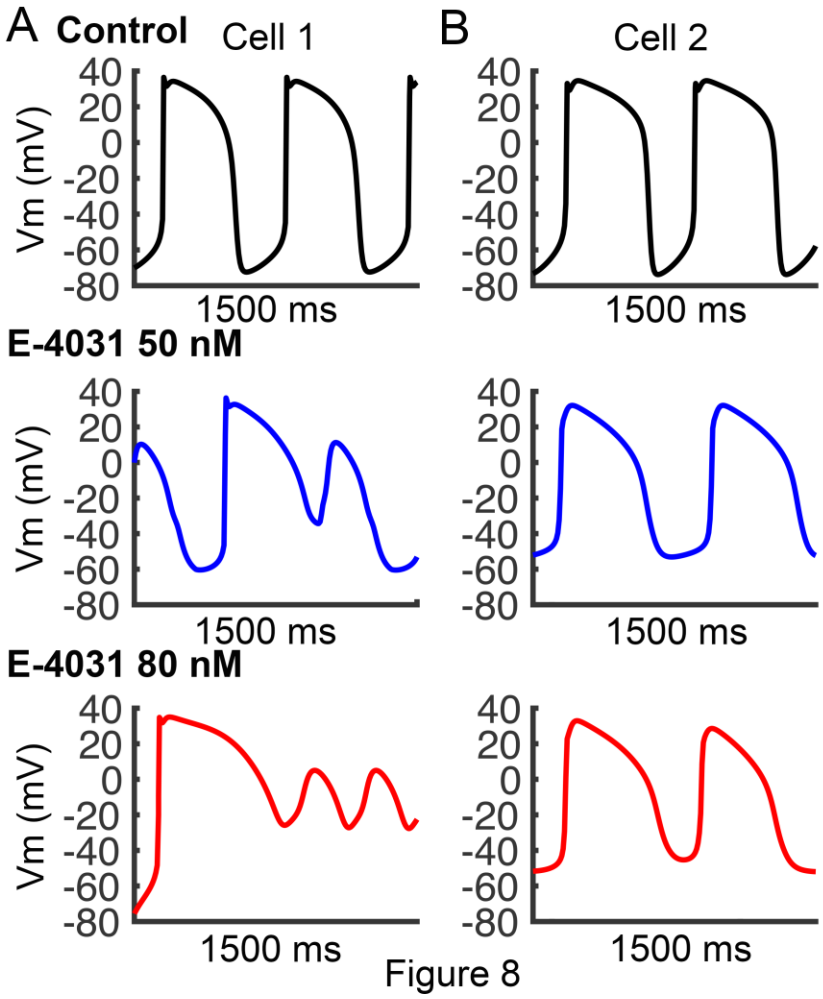
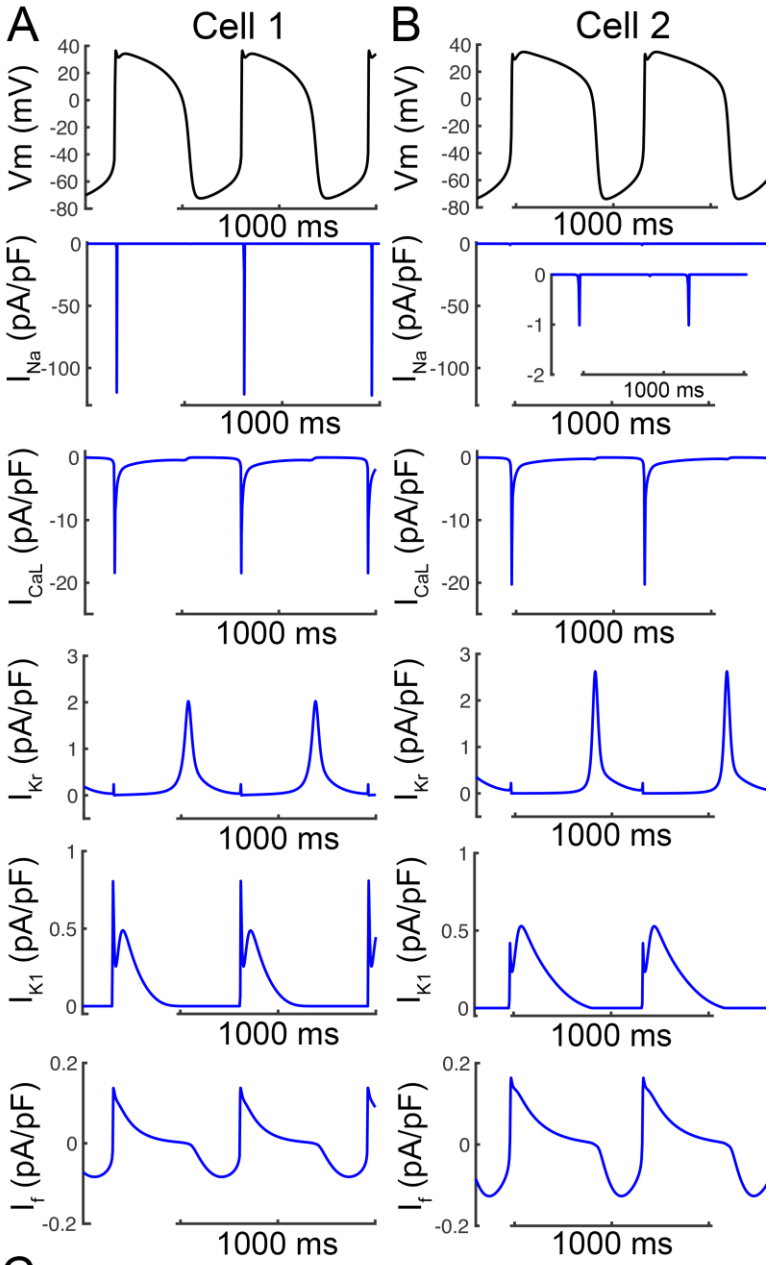


Figure 8. Variable drug response from two digital twin models of iPSC-CMs. (A) iPSC-CM model **Cell 1** (corresponding to Figure 7A). **(B)** iPSC-CM model **Cell 2** (corresponding to Figure 7B). In both panels, the **top traces** show control simulations at physiological temperature. The **middle traces** show responses in the presence of E-4031 (50 nM), and the **bottom traces** show responses with E-4031 (80 nM). Notably, **Cell 1** exhibited early afterdepolarizations (EADs) at 50 nM**,** whereas **Cell 2** developed EADs at 80 nM, demonstrating cell-specific variability relevant to precision medicine.



C

	Max Upstroke velocity (mV/ms)	Peak potential (mV)	Beating rate (bpm)
Cell 1	145.6	37.4	89.4
Cell 2	21.7	33.1	87

Figure 9

Figure 9. Electrophysiological variability in two iPSC-CM digital twins.

(A, B) Traces from simulations of **Cell 1 (**corresponding to Figure 7A**)** and **Cell 2 (**corresponding to Figure 7B**)**, generated using the parameter sets predicted by our network model. For each cell, the top panel shows the action potentials. The following

panels display the major ionic currents: fast sodium current (**I_{Na}**) L-type calcium current (**I_{caL}**), rapid delayed rectifier potassium current (**I_{Kr}**), inward rectifier potassium current (**I_{K1}**), and the hyperpolarization-activated funny current (**I_f**). Traces illustrate differences in ionic current contributions to the action potential waveform between the two cells. **(C)** Summary of max upstroke velocity, peak potential and spontaneous beating rates for iPSC-CM model **Cell 1** and **Cell 2**.