Neural network emulation of the human ventricular cardiomyocyte action potential: a tool for more efficient computation in pharmacological studies

Thomas Grandits[1,8], Christoph M. Augustin[6,7], Gundolf Haase[1], Norbert Jost[2,3], Gary R. Mirams[4], Steven A. Niederer[5], Gernot Plank[6,7], András Varró[2,3], László Virág[2], and Alexander Jung[6]

[1]Department of Mathematics and Scientific Computing, University of Graz
[2]Department of Pharmacology and Pharmacotherapy, University of Szeged
[3]ELKH-TKI, Research Group of Phamacology
[4]Centre for Mathematical Medicine & Biology, School of Mathematical Sciences, University of Nottingham
[5]Division of Imaging Sciences & Biomedical Engineering, King's College London
[6]Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging - Division of Medical Physics and Biophysics, Medical University of Graz
[7]BioTechMed-Graz
[8]NAWI Graz, University of Graz

Abstract

Computer models of the human ventricular cardiomyocyte action potential (AP) have reached a level of detail and maturity that has led to an increasing number of applications in the pharmaceutical sector. However, interfacing the models with experimental data can become a significant computational burden. To mitigate the computational burden, the present study introduces a neural network (NN) that emulates the AP for given maximum conductances of selected ion channels, pumps, and exchangers. Its applicability in pharmacological studies was tested on synthetic and experimental data. The NN emulator enabled a massive speed-up of more than 10^4 compared to regular simulations and the forward problem (find drugged AP for pharmacological parameters defined as scaling factors of control maximum conductances) on synthetic data could be solved with average root-mean-square errors (RMSE) of 0.47 mV in normal APs and of 13.6 mV in abnormal APs exhibiting early afterdepolarizations (72% of the emulated APs were alining with the abnormality, and the substantial majority of the remaining APs demonstrated pronounced proximity). This demonstrates not only very fast and mostly very accurate AP emulations but also the capability of accounting for discontinuities, a major advantage over existing emulation strategies. Furthermore, the inverse problem (find pharmacological parameters for control and drugged APs through optimization) on synthetic data could be solved with high accuracy shown by a maximum RMSE of

0.21 in the estimated pharmacological parameters. However, notable mismatches were observed between pharmacological parameters estimated from experimental data and distributions obtained from the Comprehensive in Vitro Proarrhythmia Assay initiative. This reveals larger inaccuracies which can be attributed particularly to the fact that small tissue preparations were studied while the emulator was trained on single cardiomyocyte data. Overall, our study highlights the potential of NN emulators as powerful tool for an increased efficiency in future quantitative systems pharmacology studies.

1 Introduction

Computer models of human physiology are becoming increasingly detailed and mature and the area of ventricular cardiomyocyte electrophysiology (EP) is one of the most advanced. The most updated models include fine representations of ion movements through various important channels, pumps, and exchangers, and take the complex handling of intracellular calcium accurately into account [Grandi et al., 2010, O'Hara et al., 2011, Tomek et al., 2019, Bartolucci et al., 2020]. While these models have individual strengths and limitations in replicating different aspects of physiology, pathology, and pharmacology [Corrado et al., 2020, Amuzescu et al., 2021], their degree of credibility has reached a level that has led to an increasing number of applications in academia and beyond. This holds in particular for the pharmaceutical sector, where much effort is spent on using computer modeling to reduce traditional preclinical and clinical methodologies for assessing the efficacy and safety of novel drug candidates [Mirams et al., 2011, Passini et al., 2017, Li et al., 2019, Passini et al., 2021]. To improve the regulatory assessment of a drug's proarrhythmic potential, the Comprehensive in Vitro Proarrhythmia Assay (CiPA) was proposed in 2013 following a workshop at the US Food and Drug Administration [Sager et al., 2014, Colatsky et al., 2016, Strauss et al., 2021]. A central component is a computer model of human ventricular cardiomyocyte EP that is coupled to a pharmacological model describing the interaction between a given drug and multiple arrhythmia-relevant channels [Dutta et al., 2017, Li et al., 2017, Li et al., 2019]. For a given drug, experimental channel block data are collected to inform the pharmacological model and corresponding simulations of the action potentials (AP; time course of the transmembrane potential) are performed to predict the proarrhythmic risk based on a mechanistically motivated biomarker [Chang et al., 2017, Li et al., 2017, Li et al., 2019]. The prediction is then compared with findings in experimental [Blinova et al., 2018] and clinical [Vicente et al., 2017] studies. To compute the drugged AP for given pharmacological parameters is a forward problem, while the corresponding inverse problem is to find pharmacological parameters for given control and drugged AP. Some relevant examples for the latter have been presented by [Bottino et al., 2006] who estimated pharmacological parameters from APs of canine Purkinje fibers and by Tveito *et al.* who estimated pharmacological parameters from AP biomarkers measured in human induced pluripotent stem cell-derived cardiomyocytes [Tveito et al., 2018] and several animal ventricular cardiomyocytes [Tveito et al., 2020]. Furthermore, [Jæger et al., 2021] identified the optimal polypharmacological treatment for recovering APs

of mutant ventricular cardiomyocytes based on biomarkers of simulated wild type and mutant APs.

When the models are interfaced with experimental data, attention should be paid to the inherent uncertainty in the data that results from beat-to-beat variability (intrinsic variability), cell-to-cell variability (extrinsic variability), and measurement errors (observational uncertainty) [Mirams et al., 2016]. Uncertainty propagates through the given problem from APs to estimated parameters or from parameters to predicted APs and must be properly quantified to draw reliable conclusions from the results. Multiple methodologies exist for this purpose [Oakley and O'Hagan, 2004, Mirams et al., 2016, Sher et al., 2022] but usually require many simulations, which even for ordinary differential equation (ODE)-based models of cardiomyocyte EP can become a significant computational burden when considering that each simulation includes a substantial number of beats to reach the model's limit cycle. To overcome this problem, surrogate models have emerged which approximate (emulate) chosen outputs for given inputs multiple orders of magnitude faster. In line with uncertainty quantification literature, the cardiomyocyte EP model is from now on termed the "*simulator*", whereas the surrogate model is termed the "*emulator*". Earlier work has reported on an emulator based on linear interpolation of a multi-dimensional lookup table [Mirams et al., 2014] and more recently, Gaussian process (GP) emulators have become popular. Their key advantage is that in- and outputs are modeled as random distributions which allows for rapid sampling of the posterior distributions [Chang et al., 2015, Johnstone et al., 2016, Coveney and Clayton, 2018, Ghosh et al., 2018, Rasmussen and Williams, 2019, Coveney et al., 2021] and while outputs of recently published GP emulators were relevant biomarkers of the AP [Chang et al., 2015, Johnstone et al., 2016, Coveney and Clayton, 2018, Ghosh et al., 2018, Coveney et al., 2021], the emulation of the entire AP can also be realized, e.g. through dimensionality reduction techniques such as the principle component analysis or regressing state-transition models [Mohammadi et al., 2019]. However, GP emulators are not well suited to capture discontinuities of the response surface and may thus fail to capture AP abnormalities, which is a particular drawback for pharmacological studies where bifurcations in behavior such as early afterdepolarizations (EAD) can occur [Ghosh et al., 2018]. To address this, [Ghosh et al., 2018] presented a two-step approach for the emulation of the AP duration at 90% repolarization that first sets up a GP for the location of discontinuities and then fits separate GP emulators for the output of interest either side of these boundaries. In contrast, it has been proven that neural networks (NN) can approximate even discontinuous functions with arbitrary precision in theory [Hornik et al., 1989], while recent works using NNs show empirically promising results for modeling partial differential equations containing discontinuities [Jagtap et al., 2020]. These features render NN emulators suitable emulation candidates and while [Lei and Mirams, 2021] have recently investigated NN emulation of hERG channel kinetics, [Jeong et al., 2023] proposed a neural network using AP shapes as input for the prediction of a drug's proarrhythmic risk. However, to the best of our knowledge, NN emulators have not yet been used as surrogate for cardiomyocyte EP models.

The present study introduces NN emulation of the human ventricular cardiomyocyte

AP and investigates the applicability in pharmacological studies. To this end, a NN emulator was developed based on data generated using a state-of-the-art simulator [Tomek et al., 2019, Tomek et al., 2020] and the evaluation was done for forward and inverse problems on synthetic and experimental data.

2 Methods and material

The methodology of this study including the development of the emulator and the evaluation is outlined in Figure 1.

2.1 Simulator

The simulator of [Tomek et al., 2020] (ToR-ORd-dynCl simulator) was used. This is available in *CARPentry* [Vigmond et al., 2008] and was implemented based on the published *CellML* file for the endocardial subtype[1]. Simulations were performed in *CARPentry* with the single cell tool *bench*. To compute the gating variables, the Rush-Larsen Method [Rush and Larsen, 1978] was employed, which uses an analytical solution assuming fixed voltage over a small timestep, and the remaining variables were computed by the Forward Euler method. To ensure low computational cost, we found the maximum solver and sampling time steps that still produce accurate results as follows. Various solver and sampling time steps were applied to generate APs and the biomarkers used in this study (see Appendix A) were computed and compared with those that correspond to the minimum time steps (solver: 0.005 ms; sampling: 0.01 ms). The selection criterion was a relative difference below 5% for all AP biomarkers and the chosen time steps (solver: 0.01 ms; sampling: 0.05 ms) led to a maximum relative difference of 2%. APs were stimulated at a pacing cycle length of 1000 ms using 1 ms long rectangular transmembrane current density pulses of 53 $\frac{\mu A}{cm^2}$ [Tomek et al., 2019]. To approach the simulator's limit cycle, a series of 1000 stimuli were applied for each new parameter set starting from the initial states specified in the *CellML* file (when 1000 additional stimuli were applied, the maximum intracellular [Ca^{2+}], [Cl^-], and [Na^+] changed by 1.5%, 0.15%, and 1.7%, respectively). Of note, the simulations can also be performed using the open-source software *openCARP* [Plank et al., 2021].

2.2 Emulator

2.2.1 Architecture

The emulator takes maximum conductances of channels, pumps, and exchangers as inputs and computes the corresponding AP ($V_m(t)$) after the last stimulus as output. It was trained to represent human ventricular cardiomyocytes under control and drugged conditions and the inputs were selected based on two assumptions: 1) The kinetics of channels are preserved, while the number of channels, pumps, and exchangers vary

[1] `ToRORd_dynCl_endo.cellml`, `https://github.com/jtmff/torord/tree/master/cellml`

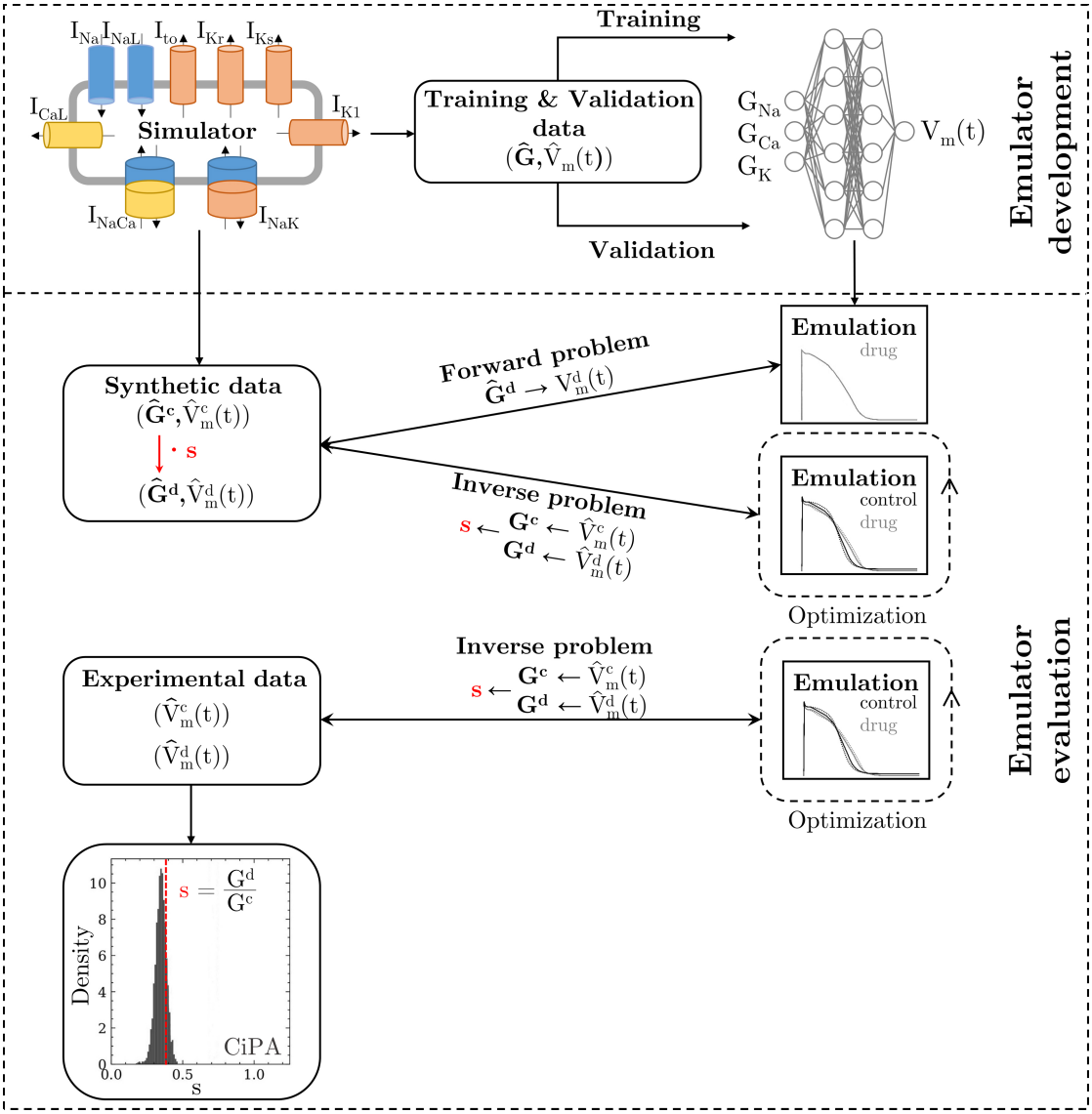

Figure 1: Methodology of this study including the emulator development and the evaluation. The simulator is visualized by a schematic human ventricular cardiomyocyte that includes all currents considered for the emulator training. Inputs of the emulator (see Figure 2) are the corresponding maximum conductances ($\mathbf{G}$) but for the sake of illustration, only three representatives are shown. Output is the AP ($V_m(t)$). Training and validation data (maximum conductances $\hat{\mathbf{G}}$ and APs $\hat{V}_{\mathrm{m}}(\mathrm{t})$) were generated by the simulator. The evaluation was performed for forward and inverse problems and to this end, the pharmacological parameter s was introduced. This describes the interaction between the drug and a given target and was defined as scaling factor of the respective maximum conductance in control conditions ($s < 1$: block, $s = 1$: no effect, $s > 1$: enhancement). Synthetic data (control maximum conductances and drugged maximum conductances obtained through scaling, and control and drugged APs) and experimental data (control and drugged APs) were used for the evaluation. The forward problem was only solved for synthetic data, whereas the inverse problem was solved for both synthetic and experimental data. When experimental data were used, estimated pharmacological parameters were compared to distributions derived from data published within the CiPA initiative [Li et al., 2017, Chang et al., 2017].

due to different expression levels [Syed et al., 2005, Groenendaal et al., 2015, Krogh-Madsen et al., 2016]. These numbers are captured in the simulator by the maximum conductance parameters (or permeability parameters but maximum conductance is used as general term here for the sake of simplicity) that determine the respective current densities; 2) Channels, pumps, and exchangers are potential (intended and unintended) drug targets and the interaction between drugs and their targets can be described by a scaling of the related maximum conductances [Brennan et al., 2009]. The corresponding scaling factors are pharmacological parameters. These assumptions allowed us to focus on maximum conductances and we considered those as inputs which either the AP is sensitive to (G_{Na}, G_{NaL}, P_{Ca}, G_{to}, G_{Kr}, G_{K1}, G_{NCX}, P_{NaK}) or which are related to common drug targets (G_{Na}, G_{NaL}, P_{Ca}, G_{to}, G_{Kr}, G_{Ks}, G_{K1}; [Crumb et al., 2016]) leading to the following selection: G_{Na}, G_{NaL}, P_{Ca}, G_{to}, G_{Kr}, G_{Ks}, G_{K1}, G_{NCX}, and P_{NaK}. AP sensitivity was quantified using a global sensitivity analysis (GSA; see Appendix B) and the inclusion threshold was a total-effect Sobol' sensitivity index (ST) above 0.1 with respect to any of the considered biomarkers (see Appendix A).

Several emulator architectures were tried on the training and validation data sets and the final choice was hand-picked as a good trade-off between high accuracy and low computational cost. We decided to utilize a two-stage emulator architecture: First, the maximum conductances $\mathbf{x}$ – normalized to the range $\mathbf{x}_i \in [-0.5,\ 0.5]$ – are encoded using a first NN (Θ_1) into a latent representation ϑ. Second, this intermediate representation parameterizes a function $f_\vartheta : \mathbb{R} \to \mathbb{R}$ defined by a second NN (Θ_2) that can be continuously evaluated to receive the emulated AP at time t. To help the second NN in computing the fast depolarization, a simple depolarization term (tanh) is added to f_ϑ. The parameters of this depolarization function are slope (d_1), offset (d_2), and amplitude (d_3), and are created by encoding the parameters through the first network, similar to the latent code. The AP approximated by the emulator is thus defined by

$$V_{\mathrm{m}}(t) \coloneqq f_\vartheta(t) + \tanh(d_1^2(t - d_2)) + d_3^2. \tag{1}$$

A schematic drawing of the emulator architecture is provided in Figure 2. Splitting the network into two parts — one for encoding the parameters into a latent space and a second one for evaluating f_ϑ — allowed us to give the emulator enough complexity without markedly increasing the computational cost: in most cases, it is desirable to compute the whole AP in the entire range, e.g. $[0,\ 1000]$ ms, and not only at a single time step. Creating a single network that computes the mapping from maximum conductances to the transmembrane potential at a single time step (compare Figure 2) is either orders of magnitudes slower than encoding the parameters into a latent vector (only done once per AP) or would require to reduce the complexity of the network, which led to inaccurate emulations in the validation. The additional depolarization term was introduced to address the difficulty of fitting the depolarization phase during training and decreased the required training time substantially. Note that the mapping from maximum conductances $\mathbf{x}$ to depolarization parameters $\{d_1, d_2, d_3\}$ is also learned through Θ_1.

The exact architecture employed – chosen by a cross-validation approach (see Sec-

tion 2.2.3) – comprised a first network (Θ_1) of four fully connected layers of 256 neurons, each to encode the parameters into the latent vector $\vartheta \in \mathbb{R}^{256}$. This first network additionally generates the parameterization for the depolarization model $\{d_1, d_2, d_3\}$. The second network, computing the APs from the latent representation (Θ_2), consisted of four fully connected layers of 64 neurons each. Exponential linear unit (ELU) activation functions [Clevert et al., 2016] were used for all layers, except for the final non-linear layer, which was modeled using a tanh activation function followed by a $[-1, 1]^{64} \to \mathbb{R}$ linear layer.

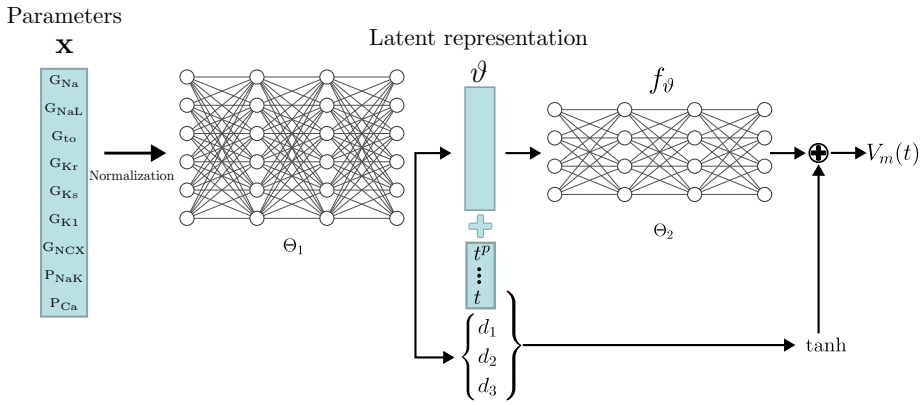

Figure 2: Conceptual architecture of the neural network emulator. The maximum conductances $\mathbf{x}$ are encoded into depolarization parameters d_i and a latent space representation ϑ that uniquely defines the time series functional $f_\vartheta : \mathbb{R} \to \mathbb{R}$. The time is normalized and encoded in polynomials up to degree 8 (t^p for $p \in \{1, 2, \ldots 8\}$), before being appended to the latent code ϑ. f_ϑ is then used in conjunction with the depolarization helper tanh to approximate the AP $V_\mathrm{m}(t)$.

2.2.2 Training

The supervised training data set X contained pairs of maximum conductance samples $\mathbf{x}$ and the corresponding AP $\hat{V}_\mathrm{m}(t)$ that was obtained from the simulator. Sobol' sequences were used to generate 40,000 maximum conductance samples, containing 20,000 maximum conductance samples between 0% and 200% of the original values and 20,000 maximum conductance samples between 50% and 150% of the original values. The first covers a range that was considered plausible in terms of physiology and pathology [Britton et al., 2017, Tomek et al., 2019], and in terms of pharmacology (where full block is plausible). The latter covers a range that was considered particularly relevant in line with experimental calibration results presented in [Tomek et al., 2019] and patch clamp measurements of channels that were exposed to 30 clinical drugs (blocks in up to the four-fold of the maximum free therapeutic concentration were analyzed in agreement with the CiPA paradigm [Li et al., 2019] [Crumb et al., 2016]. The `SALib-Sensitivity Analysis Library` [Herman and Usher, 2017] was used in the entire study for the gen-

eration of samples based on Sobol' sequences. For each maximum conductance sample, simulations were performed to obtain the corresponding 40,000 APs. APs with a transmembrane potential difference of more than 10% of the amplitude between $t = 0$ and 1000 ms were excluded. This, however, applied to only 116 APs.

Starting from the original APs, the data were first extended by 10 ms from $t \in [0, 1000]$ ms to $t \in [-10, 1000]$ ms to enable some extrapolation of V_m and hence a better alignment of the depolarization; for $t \in [-10, 0]$ ms the initial resting membrane potential $V_\mathrm{m}(0)$ was held constant. Then, the data were non-uniformly resampled to 4 kHz for $t \in [-20, -5)$ ms (resting phase) and 10 kHz for $t \in [-5, 20)$ ms (depolarization phase). The repolarization phase ($t \in [20, 1000]$ ms) was again resampled to 4 kHz.

From the initial training data set, 20% were randomly excluded to be used for validation leaving 31908 pairs of maximum conductances and corresponding APs for training. Figure 3 shows processed APs that were used for training and validation. For better visualization, only $\approx 25\%$ are included but this already illustrates the large variety of APs spanning from normal to abnormal APs.

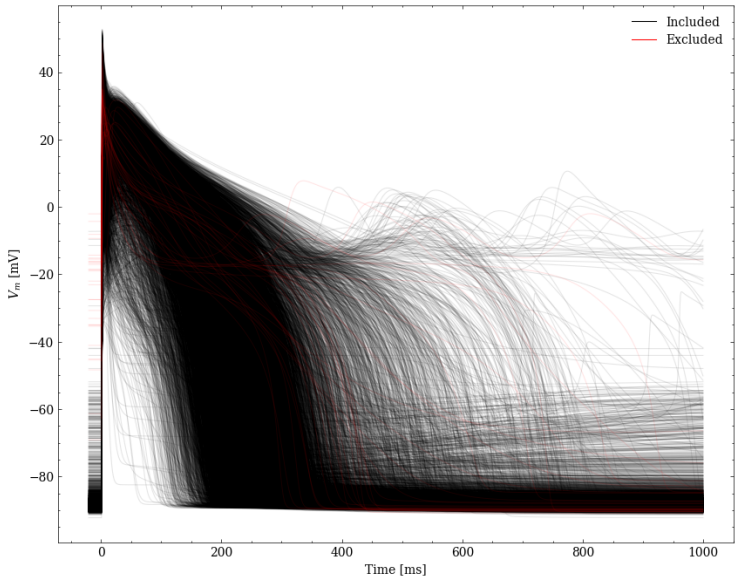

Figure 3: Processed APs used for training and validation. For better visualization, only 4000 of the set are shown.

Although different regularization schemes such as variational losses (e.g. $\frac{\partial V_\mathrm{m}}{\partial t}$) were tried, the wealth of training data allowed us to define the training loss purely in terms of mean-squared-error (MSE)

$$\mathcal{L}(\Theta_1, \Theta_2) = \frac{1}{2|\mathbb{T}|} \sum_{(\mathbf{x}, \hat{V}_\mathrm{m}) \in \bar{X} \subset X} \sum_{t \in \mathbb{T}} \left(f_{\vartheta_{\Theta_1}(\mathbf{x}), \Theta_2}(t) - \hat{V}_\mathrm{m}(t) \right)^2 \mathrm{d}t, \tag{2}$$

where $f_{\vartheta_{\Theta_1}(\mathbf{x}), \Theta_2}$ describes the output of the emulator using the current NN weights Θ_1/Θ_2, $\bar{X}$ refers to the current training batch and X is the training data set containing

both target APs $\hat{V}_\mathrm{m}$and corresponding maximum conductances $\mathbf{x}$. For the training, increasing batch sizes ($|\bar{X}|$ 1250 to 1800) were used, both in terms of the entire AP and considering only subsets over time $T \subset \mathbb{T}$ from $\frac{1}{16}|\mathbb{T}|$ to $|\mathbb{T}|$, resulting in a reduction of training time needed. The neural network was trained for a total number of 5000 epochs using the first-order gradient-based algorithm ADAM [Kingma and Ba, 2017]. The training time was approximately 4 h on a Nvidia A100 (Nvidia Corporation, USA).

The simulation of 1000 APs (see Section 2.1) sampled at a resolution of 20 kHz took 293 s on one core of a AMD Ryzen Threadripper 2990WX (clock rate: 3.0 GHz). The emulation of an AP (that approached the limit cycle after 1000 stimuli) sampled at a resolution of 20 kHz for $t \in [-10,\ 1000]$ ms took 18.7 ms on a AMD Ryzen 7 3800X (clock rate: 4.5 GHz) and 1.59 ms on a Nvidia GeForce GTX 1660 Super GPU, including synchronization and data copy overhead. This is a speed-up of $1.57 \cdot 10^4$ and $1.84 \cdot 10^5$, respectively.

2.2.3 Validation

A cross-validation approach was used to quantify and compare the performance of various emulator architectures. The validation was based on 20% of the initial training data set (7976 pairs of maximum conductances and corresponding APs, see Section 2.2.2). For each pair, the emulated AP $V_\mathrm{m}$was compared against the simulated AP $\hat{V}_\mathrm{m}$, given the same maximum conductances. The mismatch was quantified by the root-mean-squared error (RMSE) defined as

$$\mathrm{RMSE}(V_\mathrm{m}, \tilde{V}_m) \coloneqq \sqrt{\frac{1}{|T|}\int_T \left(V_\mathrm{m}(t) - \hat{V}_\mathrm{m}(t)\right)^2 \mathrm{d}t}.$$

The mismatch was also quantified in terms of AP biomarkers $\mathbf{b} \in \mathbb{R}^N$ (see Appendix A) and normalized maximum conductances $\mathbf{x}$ (see Section 2.2.4) in which case the RMSE was defined as

$$\mathrm{RMSE}(\mathbf{b}, \tilde{\mathbf{b}}) \coloneqq \sqrt{\frac{1}{N}||\mathbf{b} - \tilde{\mathbf{b}}||_2^2} \quad \text{and} \quad \mathrm{RMSE}(\mathbf{x}, \tilde{\mathbf{x}}) \coloneqq \sqrt{\frac{1}{N}||\mathbf{x} - \tilde{\mathbf{x}}||_2^2},$$

for N samples.

2.2.4 Time series fitting and estimation of maximum conductances and pharmacological parameters

Time series fitting is the basis for solving the inverse problem. To fit a given AP $\hat{V}_\mathrm{m}$, defined on a subset of the trained domain $\hat{\mathbb{T}} \subset \mathbb{T}$, the first step was to choose a trial set of maximum conductances $\mathbf{x}_0$. Then, for the given trial set, the corresponding AP was emulated and the trial set of maximum conductances was iteratively updated by solving the following minimization problem:

$$\min_{\mathbf{x}, t_0} \frac{1}{2|\hat{\mathbb{T}}|}\sum_{t \in \hat{\mathbb{T}}} \left(V_\mathrm{m}(\mathbf{x}, t - t_0) - \hat{V_\mathrm{m}}(t)\right)^2 \mathrm{d}t + \frac{\lambda_{\mathbf{x}_0}}{2}||\mathbf{x} - \hat{\mathbf{x}}||_2^2 + \delta_{[-0.5,0.5]}(\mathbf{x}), \tag{3}$$

where $V_{\mathrm{m}}(\mathbf{x}, t) \coloneqq f_{\vartheta_{\Theta_1}(\mathbf{x}),\Theta_2}(t)$ is a shorthand for the emulator approximation function and t_0 is a temporal offset parameter helping in fitting the exact depolarization timing. Here, $\delta_{[-0.5,0.5]}(\mathbf{x})$ is the element-wise indicator function on the normalized feasible parameter space $[-0.5, 0.5]$. The minimization was done using ADAM [Kingma and Ba, 2017] combined with a projection on the feasible space.

To estimate maximum conductances for a given control AP, the control AP was fitted using the original maximum conductance values as initial guesses and priors: $\mathbf{x}_0 = \hat{\mathbf{x}} = \mathbf{0}$. To estimate maximum conductances for a given drugged AP, the drugged AP was fitted using the maximum conductances estimated for the respective control AP as initial guesses and priors: $\mathbf{x}_0 = \hat{\mathbf{x}} = \mathbf{x}^c$. The pharmacological parameters (scaling factors of control maximum conductances) were computed as element-wise ratios between the drugged and control maximum conductances $(\mathbf{s})_i \coloneqq \frac{(\mathbf{G}^d)_i}{(\mathbf{G}^c)_i}$ but here, $\mathbf{G}^d$ and $\mathbf{G}^c$ are the non-normalized maximum conductances, where $(\mathbf{G}^c)_i > 0$.

Since multiple sets of maximum conductances produced similarly good fits of the given AP, the parameter $\lambda_{\mathbf{x}_0}$was introduced which minimizes the difference between original and control maximum conductances and between control and drugged maximum conductances, respectively. The value was chosen to be $\lambda_{\mathbf{x}_0} = 10$ with respect to the highest accuracy found for the synthetic data set #1 that was generated for the evaluation (see Section 2.3.1).

2.3 Evaluation

The evaluation was performed for forward and inverse problems in pharmacological studies on synthetic and experimental data.

2.3.1 Data

Synthetic data Two sets of synthetic data were created using the simulator and each of the sets consisted of control and drug data with pairs of maximum conductances and corresponding APs.

The control data were the same in both sets. They were created using an experimentally calibrated population of 100 synthetic cardiomyocytes [Britton et al., 2013, Muszkiewicz et al., 2016]. and to this end, Sobol' sequences were used to generate samples of maximum conductances with values between 50% and 150% of the original values. Maximum conductance samples that produced APs without abnormalities (checked for the last two consecutive APs; see Appendix A), and with seven biomarker values (derived from the last AP; see Appendix A) in agreement with experimental ranges (Table 1) were included in the population. Please note that the experimental ranges were not derived from the data set described in Section 2.3.1.

Data set #1: The motivation for creating data set #1 was to evaluate the emulator on data of normal APs.

Drug data were created using 100 synthetic drugs represented by a set of pharmacological parameters. Each synthetic drug was built to have four different targets, with all channels, pumps, and exchangers related to the emulator inputs considered as potential targets. To this end, 100 samples of four pharmacological parameters, each with values between 0.5 (50% block) and 1.5 (50% enhancement) were randomly generated. The synthetic drugs were applied to the entire synthetic cardiomyocyte population by scaling each of the relevant control maximum conductances with the corresponding pharmacological parameter. The samples that produced APs without abnormalities (checked for the last two consecutive APs; see Appendix A) were included in the data set. No sample was excluded and thus, the data set consisted of 100 control data pairs and 10,000 drug data pairs.

Data set #2: The motivation for creating data set #2 was to test the emulator on data of abnormal APs showing the repolarization abnormality EAD. This is considered a particularly relevant AP abnormality in pharmacological studies because of their role in the genesis of drug-induced ventricular arrhythmia's [Weiss et al., 2010]. Drug data were created using ten synthetic drugs with the hERG channel and the Cav1.2 channel as targets. To this end, ten samples with pharmacological parameters for G_{Kr} and P_{Ca} (Table 2) were generated and the synthetic drugs were applied to the entire synthetic cardiomyocyte population by scaling G_{Kr} and P_{Ca} with the corresponding pharmacological parameter. The samples that produced APs with EADs (see Appendix A) but a transmembrane potential difference of not more than 10% of the amplitude between $t = 0$ and 1000 ms (checked for the last AP) were included. The data set consisted of 100 control data pairs and 122 drug data pairs.

Table 1: AP biomarkers and their experimental ranges used to generate the population of synthetic cardiomyocytes. These were adopted from [Passini et al., 2017]. Experimental data were collected at 37 °C in small right ventricular trabeculae and papillary tissue preparations obtained from healthy human hearts during pacing with a cycle length of 1000 ms [Britton et al., 2017].

AP biomarker	Unit	Min	Max
RMP	mV	−95	−80
$dVmMax$	mVms^{-1}	100	1000
$Peak$	mV	10	55
APD_{40}	ms	85	320
APD_{50}	ms	110	350
APD_{90}	ms	180	440
Tri_{90-40}	ms	50	150

Experimental data APs were recorded in small right ventricular trabeculae and papillary tissue preparations that were obtained from healthy human hearts [Orvos et al.,

Table 2: Pharmacological parameter samples (synthetic drugs) with scaling factors for G_{Kr} and P_{Ca} to generate the drug data of data set #2.

ID	1	2	3	4	5	6	7	8	9	10
G_{Kr}	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14
P_{Ca}	1.20	1.22	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38

2019]. The conventional microelectrode technique was used for recording at a temperature of 37 °C. Stimulation of APs was done at a pacing cycle length of 1000 ms using a pair of platinum electrodes that provided rectangular current pulses of 2 ms duration. To allow the preparations to equilibrate, stimuli were delivered for at least 60 min before the measurements started. Measurements were performed under control conditions and after administration of five channel-blocking drugs at one concentration in multiple preparations. Drugs were cisapride (30 nM), dofetilide (10 nM), sotalol (30 µM), terfenadine (1 µM), and verapamil (300 nM). The last ten consecutive APs of each measurement were analyzed to quantify the beat-to-beat variability. Overall, the beat-to-beat variability was found to be small (standard deviation in all APs below 7 mV before the peak due to time alignment mismatch and below 2 mV after the peak) and thus, the last ten consecutive APs of each measurement were averaged. In most of the preparations, the standard deviation between beats did not vary over time and thus, no temporal correlation of noise was assumed. Averaging also reduced the noise level. The experimental data set contained one pair of averaged control and drugged AP per preparation per drug. Pairs were excluded if the biomarker values (see Appendix A) of the control or the drugged AP were not in the range found in the training data (see Section 2.2.2). This applied to seven pairs and the final data set contained three pairs for cisapride, dofetilide, sotalol, and terfenadine, and one pair for verapamil. All measurements were performed at the University of Szeged, Hungary, and conformed to the principles of the Declaration of Helsinki. Protocols were approved by the National Scientific and Research Ethical Review Boards (4991-0/2010-1018EKU [339/PI/010]).

The raw data were obtained without filtering but some filtering was applied before interfacing the data with the emulator. The APs contained a stimulus artifact between 0 and 1.5 ms that was filtered as follows. For each of the last ten consecutive APs, the transmembrane potential V_m closest to the end of the recorded time series was defined as the resting transmembrane potential RMP and $V_m(t) = RMP$ was set for $t \in [0, 1.75]$ ms. Then, the APs were resampled at 100 kHz and a low-pass filtering was performed with a second-order butterworth filter (cutoff at 2.5 kHz) to reduce the high-frequency noise of the signal. Finally, the filtered APs were averaged and the averaged AP was again resampled at 1 kHz for $t \in [15, 1000]$ ms (repolarization) and 100 kHz in $t \in [0, 15)$ ms (depolarization). An example comparison of a raw and a filtered averaged AP is given in Figure 4.

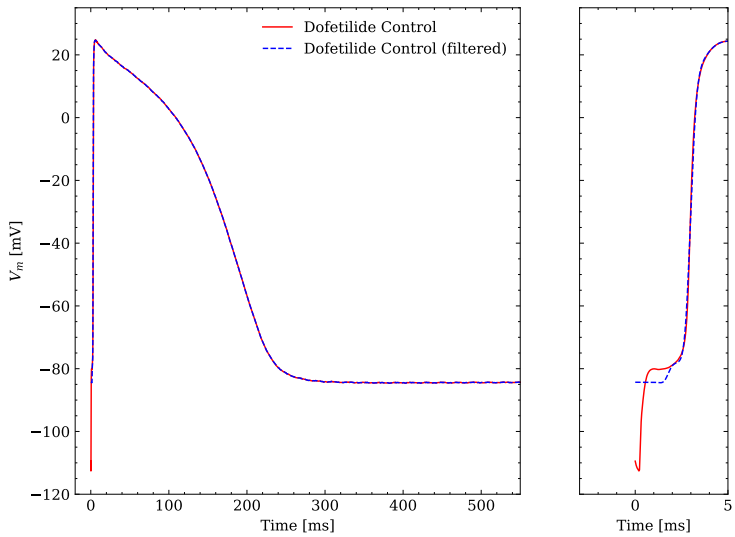

Figure 4: Comparison of an averaged raw and an averaged filtered experimental AP. One dofetilide control AP is shown as example.

2.3.2 Forward problem

The emulator evaluation for the forward problem, i.e. to find the drugged AP for given pharmacological parameters, was only performed on synthetic data since maximum conductances were not available experimentally. The data sets #1 and #2 were used to consider data with normal APs and with abnormal APs exhibiting EADs. Pharmacological parameters are not inputs of the emulator but drugged maximum conductances that were computed as control maximum conductances scaled by the given pharmacological parameters (see Section 2.3.1). These were used to emulate drugged APs. The RMSE was used to quantify the mismatch between the emulated and the ground truth AP.

2.3.3 Inverse problem

The emulator evaluation for the inverse problem, i.e. to find the pharmacological parameters for given control and drugged APs (through optimization), was performed on both synthetic and experimental data. When using synthetic data, the data set #1 was used including data with normal APs. First, control and drugged maximum conductances were estimated based on control and drugged APs and then, pharmacological parameters were computed as ratios of drugged and control conductances (see Section 2.2.4). The mismatch between estimated and ground truth maximum conductances were quantified using the error that is defined through

$$\mathbf{e}_n := (\mathbf{x} - \hat{\mathbf{x}}) \,, \tag{4}$$

where $\mathbf{x}$ and $\hat{\mathbf{x}}$ are the normalized estimated and ground truth maximum conductances (see Section 2.2.1). Similarly, we computed the mismatch between estimated and ground

truth scaling factor vectors ($\mathbf{s}$ and $\hat{\mathbf{s}}$ respectively) as

$$\mathbf{e}_s \coloneqq (\mathbf{s} - \hat{\mathbf{s}})\,. \tag{5}$$

When using experimental data, maximum conductances and pharmacological parameters were estimated in the same way but due to a lack of experimental maximum conductances, the mismatch between estimated and ground truth values could not be quantified. Instead, the estimated pharmacological parameters were compared with distributions computed from data published within the CiPA initiative [Li et al., 2017, Chang et al., 2017] (CiPA distributions). The data set [2] includes 2000 IC_{50} values and Hill coefficients for each drug and for up to seven targets (I_{Na}, I_{NaL}, I_{CaL}, I_{to}, I_{Kr}, I_{Ks}, I_{K1}). The pore-block model [Brennan et al., 2009] was used to obtain the corresponding scaling factors.

3 Results

3.1 Evaluation

3.1.1 Forward problem

The emulator evaluation for the forward problem was only done on synthetic data and both data sets #1 and #2 (see Section 2.3.1) were used to analyze the solution accuracy for normal and abnormal APs exhibiting EADs.

The data set #1 was used first and Figure 5 illustrates the distribution of RMSEs between emulated and ground truth drugged APs. The average RMSE over all APs was 0.47 mV and only for a few APs the RMSE was >1 mV with 1.5 mV being the maximum. Largest mismatches were located in the phases 0 and 3 of the AP. While the mismatches in phase 3 were simply a result of imperfect emulation, the mismatches in phase 0 were a result of the difficulty in matching the depolarization time exactly.

Figure 6 shows the distribution of biomarker mismatches between emulated and ground truth drugged APs. The low RMSEs between the APs translated into low RMSEs between the AP biomarkers.

The data set #2 was used second and Appendix C shows all emulated APs. In total, 72% of the APs exhibited EADs (classified based on the criterion outlined in Appendix A) in agreement with the ground truth drugged APs, and the substantial majority of the remaining APs showcase a notable proximity to the threshold of manifesting EADs. Figure 7 illustrates the distribution of RMSEs between emulated and ground truth drugged APs. The average RMSE over all APs was 13.6 mV with 43.2 mV being the maximum. Largest mismatches were located in phase 3 of the AP, in particular in emulated APs that did not fully repolarize.

[2] `https://github.com/FDA/CiPA/tree/master/Hill_fitting/results`

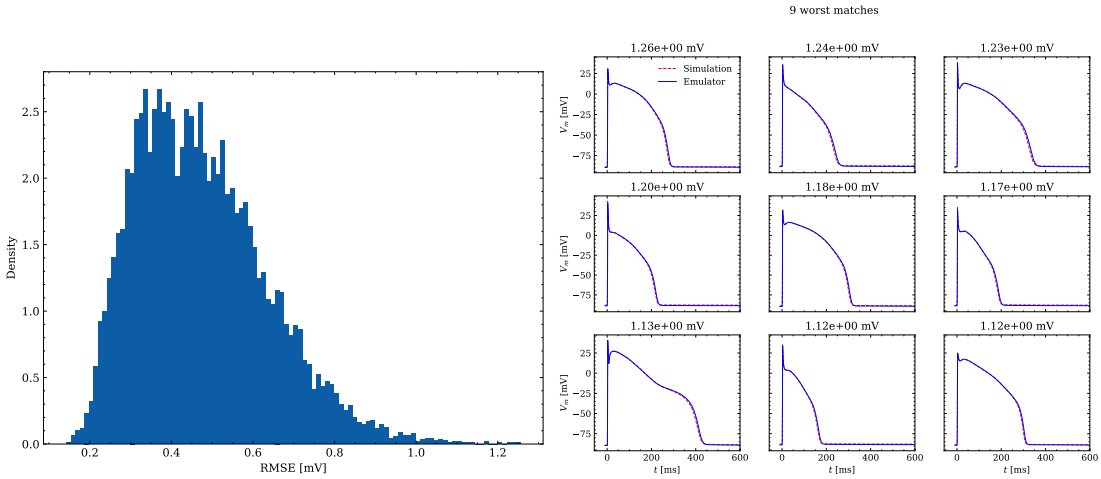

Figure 5: Analysis of solution accuracy of the forward problem on synthetic data including normal APs (drug data of data set #1). Left: histogram of RMSEs for the APs, right: APs with the largest RMSEs. The RMSE is given above each subplot.

3.1.2 Inverse problem based on synthetic data

The emulator evaluation for the inverse problem was first done using synthetic data (data set #1, see Section 2.3.1). Control and drugged APs could be fitted with an average RMSE of 0.8 mV. Largest mismatches were located in phase 0 and 3 of the AP for the reasons given above (see Section 3.1.1). Figure 8 shows the distribution of the errors between the estimated and the ground truth maximum conductances and pharmacological parameters. For both the maximum conductances (RMSE ≤ 0.18) and the related pharmacological parameters (RMSE ≤ 0.21) the errors were closely distributed around zero. However, the RMSEs increased from the control maximum conductance over the drugged maximum conductance to the pharmacological parameters and there were distinctive differences among maximum conductances and related pharmacological parameters with the smallest for G_{Kr} and the largest for G_{Ks}

3.1.3 Inverse problem based on experimental data

The emulator evaluation for the inverse problem was then done on experimental data (Section 2.3.1).

Figure 9 shows the fitted and the ground truth APs for all drugs. Control and drugged APs could be fitted with average RMSEs shown in Table 3. The largest mismatch was located in phase 0 for most APs, which was the result from an imperfect matching of the exact depolarization timing.

Figure 10 compares the estimated pharmacological parameters and the CiPA distributions. Pharmacological parameters that fell into the range spanned by $\mu \pm (0.15 + \sigma)$ of the CiPA distribution, where μ, σ are the distribution's mean and standard deviation respectively, were classified as successfully estimated while the others were classified as

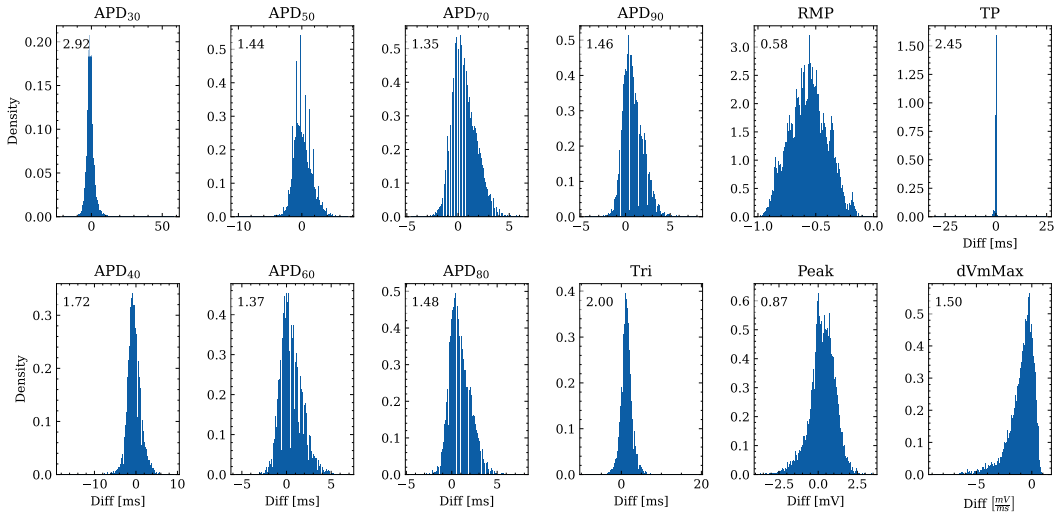

Figure 6: Analysis of solution accuracy of the forward problem on synthetic data including normal APs (data set #1) with respect to AP biomarkers. Histograms of mismatches for each biomarker are shown and the RMSE is given in the upper left corner. Note that potential staircasing artifacts in the APDs are related to the temporal discretization of the APs generated from the simulator.

unsuccessfully estimated. In total, 50% of the pharmacological parameters could be estimated succesfully and while all pharmacological parameters related to G_{Ks}could be successfully estimated, unsuccessfully estimated parameters were found across all maximum conductances, in particular related to $G_{K1}$for which all pharmacological parameters could not be successfully estimated (Table 4).

4 Discussion

NN emulation of the human ventricular cardiomyocyte AP was introduced and the applicability in pharmacological studies was investigated.

Table 3: Average RMSE over control and drugged APs measured in all preparations per drug. All values in mV.

Drug	RMSE Control	RMSE Drug
Cisapride	1.53	2.00
Dofetilide	2.05	1.73
Sotalol	1.40	2.51
Terfenadine	1.22	1.08
Verapamil	1.93	2.21

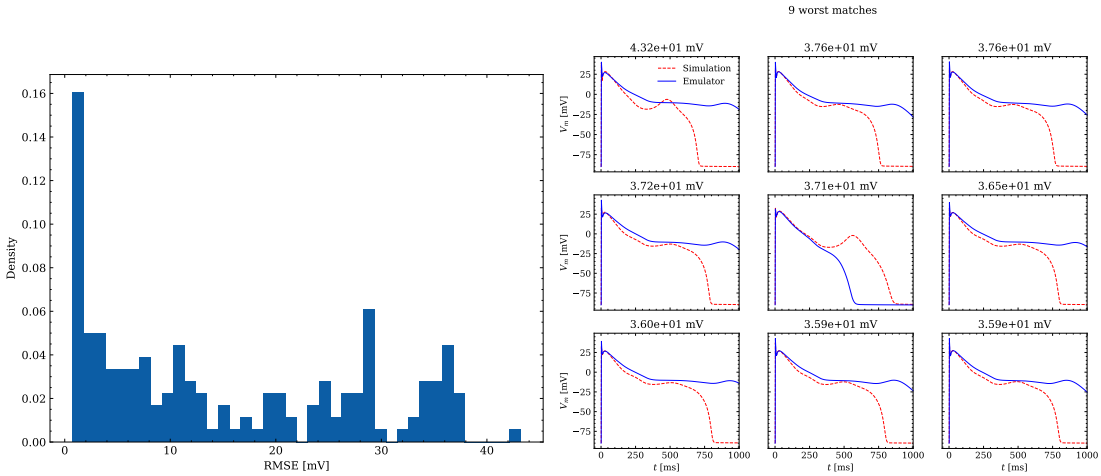

Figure 7: Analysis of solution accuracy of the forward problem on synthetic data including abnormal APs exhibiting EADs (drug data of data set #2). Left: histogram of RMSEs for the APs, right: APs with the largest RMSEs. 123 of the 170 emulated traces exhibit the expected EADs (based on the criterion outlined in Appendix A). The RMSE is given above each subplot. All emulated APs are shown in Appendix C

Table 4: Pharmacological parameters related to maximum conductances that were considered successfully or unsuccessfully estimated across all preparations and drugs

	G_{Na}	G_{NaL}	G_{to}	G_{Kr}	G_{Ks}	G_{K1}	P_{Ca}	Total
Successful	6	6	8	3	9	0	5	37
Unsuccessful	3	1	5	10	0	10	8	37
Ratio	0.67	0.86	0.62	0.23	1.00	0.00	0.38	0.50

4.1 Evaluation

The presented NN emulator enables a massive speed-up compared to regular simulations ($> 10^4$) and the evaluation for the forward problem on synthetic data showed also highly accurate AP emulations. High accuracy was found in normal APs (average RMSE was 0.47 mV; Figure 5) and mostly also in abnormal APs exhibiting EADs (of the emulated APs, 72% exhibited alignment with the abnormality, and the substantial majority of the remaining APs demonstrated pronounced proximity, while the average RMSE was 13.6 mV; Figure 7). Increasing the amount of training data within the relevant range could lead to further enhancements in accuracy for abnormal APs. Nevertheless, this observation demonstrates that the emulator is also capable of accounting for discontinuities of the response surface. This is particularly useful in pharmacological studies and a key advantage over existing emulation approaches [Chang et al., 2015, Johnstone et al., 2016, Coveney and Clayton, 2018, Ghosh et al., 2018, Rasmussen and Williams,

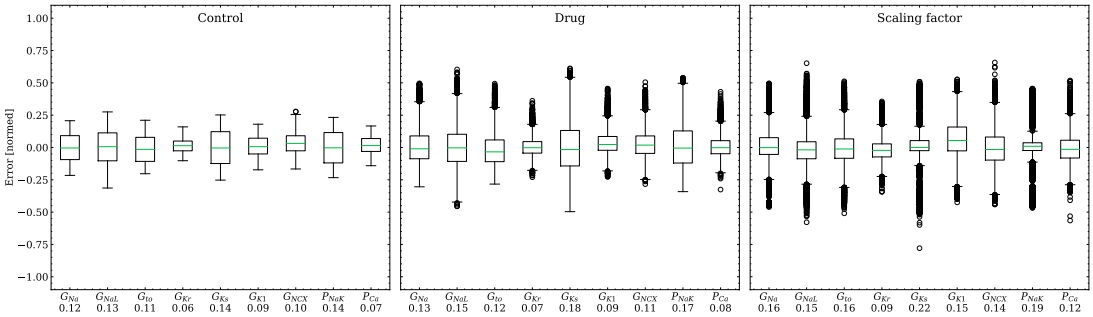

Figure 8: Analysis of solution accuracy of the inverse problem on synthetic data (data set #1). Left: boxplot of errors between normalized estimated and ground truth control maximum conductances, middle: boxplot of errors between normalized estimated and ground truth drugged maximum conductances, right: boxplot of errors between estimated and ground truth pharmacological parameters. Error definitions are given in (4) and (5). The RMSE over all data is given below each parameter.

2019, Coveney et al., 2021].

The emulator was further evaluated for the inverse problem on synthetic and also on experimental data. Maximum conductances and related pharmacological parameters could be widely estimated with high accuracy on synthetic data (RMSE ≤ 0.18 and ≤ 0.21 for all maximum conductances and pharmacological parameters, respectively; Figure 8).

The RMSEs increased from the control maximum conductance over the drugged maximum conductance to the pharmacological parameters which may be because the estimation of drugged maximum conductances depends on the control maximum conductances and the computation of pharmacological parameters depends on both control and and drugged maximum conductances (see Section 2.2.4) allowing that errors can propagate and amplify. Distinctive differences were observed among the maximum conductances and related pharmacological parameters and the largest RMSEs were found for G_{Ks} throughout. This can be attributed to various degrees of parameter identifiability [Sarkar and Sobie, 2010, Zaniboni et al., 2010, Groenendaal et al., 2015, Jæger et al., 2019, Whittaker et al., 2020] and the results agree with the GSA that indicates almost non-identifiability of G_{Ks} (Appendix B).

Larger inaccuracies were found in the inverse problem solutions on experimental data (Figure 10, Table 3). The first reason may be low parameter identifiability and we want to highlight inaccuracies in estimating the pharmacological parameters related to G_{Kr}, P_{Ca}, and G_{NaL} when the hERG channel was blocked in parallel to the Cav1.2 channel (verapamil) or in parallel to both the Cav1.2 and the Nav1.5-late channel (terfenadine). The hERG channel block (prolongation of the AP), and the Cav1.2 and Nav1.5-late channel block (shortening of the AP) are known to have opposite effects on the AP [Orvos et al., 2019]. At the given drug concentrations, these effects were

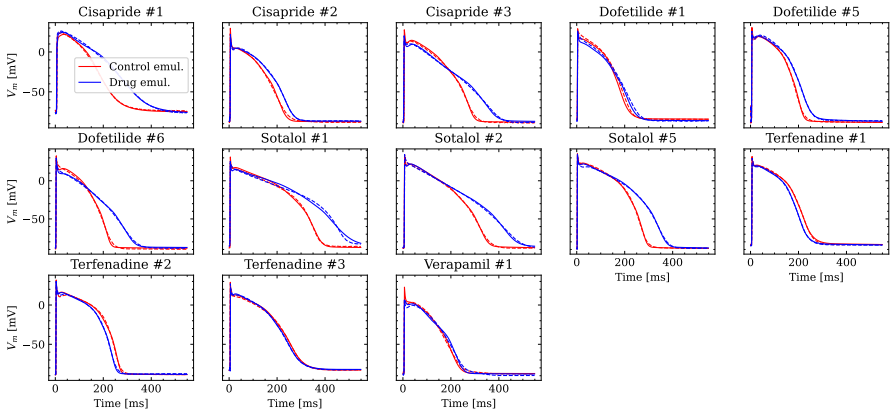

Figure 9: Analysis of fit quality of the inverse problem on experimental data. Comparison of the fitted APs (solid lines) and the experimental APs (dashed lines) at control (red) and after drug administration (blue) for all drugs.

apparently counterbalancing, which resulted in negligible changes of the AP (Figure 9). This situation made the estimation of pharmacological parameters very challenging and led to particularly large inaccuracies for terfenadine. The accurate estimations of the pharmacological parameters related to G_{Ks} are surprising at first in light of the almost non-identifiability. This was due to the combination of two factors: 1) different from synthetic data, the drugs at the given concentrations did not affect the corresponding KCNQ1-MinK channel and 2) the difference between control and drugged maximum conductance was weakly enforced to be minimal (see Section 2.2.4) which leads to almost no difference in non-identifiable parameters and hence, to a pharmacological parameter of one. The accuracy will likely be much lower in drugs that affect the KCNQ1-MinK channel.

The second and probably main reason for the inaccuracies may be the fact that the data were collected in small tissue preparations, whereas the emulator was trained on data generated by a simulator that represents single cardiomyocytes. APs in small tissue preparations are slightly different from those in single cardiomyocytes. Differences can arise from electrotonic coupling and the mixture of cells including fibroblasts that are able to modify the EP [Kohl and Gourdie, 2014, Mayer et al., 2017, Hall et al., 2021]. This can hamper the fitting of the APs and consequently, the estimation of the maximum conductances and pharmacological parameters.

4.2 Emulator

In addition to the aforementioned strengths of the presented NN emulator, some further valuable features are worth mentioning that arise from the continuous nature of the emulator. First, the AP can be emulated and fitted at any desired resolution. Second, timing offsets, e.g. between stimuli in the data to be fitted and the training data, can be accounted for using t_0 in (3) without retraining. Last but not least, the transmem-

brane potential gradient $\frac{\partial V_{\mathrm{m}}}{\partial t}$, recently highlighted in terms of proarrhythmic potential prediction [Jeong et al., 2022], can be continuously derived and is not dependent on the temporal discretization.

4.3 Limitations and future work

Some limitations have to be considered. First, the emulator has only maximum conductances as inputs. Although these explain much of the AP variability seen between cardiomyocytes [Britton et al., 2013, Muszkiewicz et al., 2016], the inclusion of parameters related to the channel kinetics might enable a more detailed consideration of drug effects in pharmacological studies. Channel kinetics determine the contribution of the corresponding current to the AP generation in different phases and can thus also modulate drug effects. However, the expansion of the input space might be unsuitable for solving the inverse problem when only AP data are used. Second, the interaction between drugs and their targets is solely captured through scaling of the related maximum conductance at control, which is mostly adequate but in fact an oversimplification [Brennan et al., 2009]. The interaction can be dependent on time, voltage, and channel state, which requires the use of Markov models with multiple more parameters [Brennan et al., 2009, Li et al., 2017]. Again, this expands the input space and might be unsuitable for solving the inverse problem when only AP data are used. Moreover, drugs that are applied over a longer period of time can also cause modifications of the maximum conductances through changes in gene expression [Shim et al., 2023]. This requires attention to avoid misinterpretations of found blocking or enhancement effects, e.g. by estimating the control maximum conductances again after a washout procedure. Third, the inverse problem was only solved for AP data obtained from one single stimulation protocol. [Johnstone et al., 2016] have shown that the usage of AP data obtained from various stimulation protocols can improve the parameter identifiability and thus, the accuracy of parameter estimates. To be able to use those data in the presented approach, the pacing cycle length must be included as additional input in the emulator and the emulator may be trained on more than the last AP of the pacing series. This would also allow to capture alternans. Last but not least, the number of drugs and concentrations considered in the inverse problem on experimental data poses a limitation. The ultimate goal is to have a tool that provides highly accurate solutions for drugs with different targets and concentrations. To this end, analyses must be extended by data obtained from a series of available and well characterized drugs. The data should be collected in single cardiomyocytes in order to minimize the discussed inaccuracies that stem from the use of tissue preparation data. This should be addressed in future work. Of note, the presented approach can also be straightforwardly applied to other transients, e.g. of the intracellular $[\mathrm{Ca}^{2+}]$ or of the sarcomere length.

4.4 Conclusion

This paper introduced NN emulation of the human ventricular cardiomyocyte AP and tested its applicability in pharmacological studies. The computational cost of the NN

emulator was compared to that of the simulator, revealing a massive speed-up of more than 10^4. The accuracy of solving the forward problem on synthetic data was found to be high for normal APs and this hold mostly true for abnormal APs exhibiting EADs. This advantage distinguishes our novel approach from existing emulation methods. While larger inaccuracies were observed when utilizing experimental data – a limitation thoroughly discussed and particularly inherent to the fact that small tissue preparations were studied while the emulator was trained on single cardiomyocyte data – the accuracy of solving the inverse problem on synthetic data remained high. Collectively, these findings underscore the potential of NN emulators in improving the efficiency of future quantitative systems pharmacology studies.

5 Code availability

The trained emulator is available as a python package under the link `https://github.com/thomgrand/cardiomyocyte_emulator`[3].

References

[Amuzescu et al., 2021] Amuzescu, B., Airini, R., Epureanu, F. B., Mann, S. A., Knott, T., and Radu, B. M. (2021). Evolution of mathematical models of cardiomyocyte electrophysiology. *Mathematical Biosciences*, 334:108567.

[Bartolucci et al., 2020] Bartolucci, C., Passini, E., Hyttinen, J., Paci, M., and Severi, S. (2020). Simulation of the effects of extracellular calcium changes leads to a novel computational model of human ventricular action potential with a revised calcium handling. *Frontiers in Physiology*, 11.

[Blinova et al., 2018] Blinova, K., Dang, Q., Millard, D., Smith, G., Pierson, J., Guo, L., Brock, M., Lu, H. R., Kraushaar, U., Zeng, H., Shi, H., Zhang, X., Sawada, K., Osada, T., Kanda, Y., Sekino, Y., Pang, L., Feaster, T. K., Kettenhofen, R., Stockbridge, N., Strauss, D. G., and Gintant, G. (2018). International multisite study of human-induced pluripotent stem cell-derived cardiomyocytes for drug proarrhythmic potential assessment. *Cell Reports*, 24(13):3582–3592.

[Bottino et al., 2006] Bottino, D., Penland, R. C., Stamps, A., Traebert, M., Dumotier, B., Georgieva, A., Helmlinger, G., and Lett, G. S. (2006). Preclinical cardiac safety assessment of pharmaceutical compounds using an integrated systems-based computer model of the heart. *Progress in Biophysics and Molecular Biology*, 90(1-3):414–443.

[Brennan et al., 2009] Brennan, T., Fink, M., and Rodriguez, B. (2009). Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *European Journal of Pharmaceutical Sciences*, 36(1):62–77.

[3]The code is licensed under AGPLv3, see `https://www.gnu.org/licenses/agpl-3.0.en.html` for details

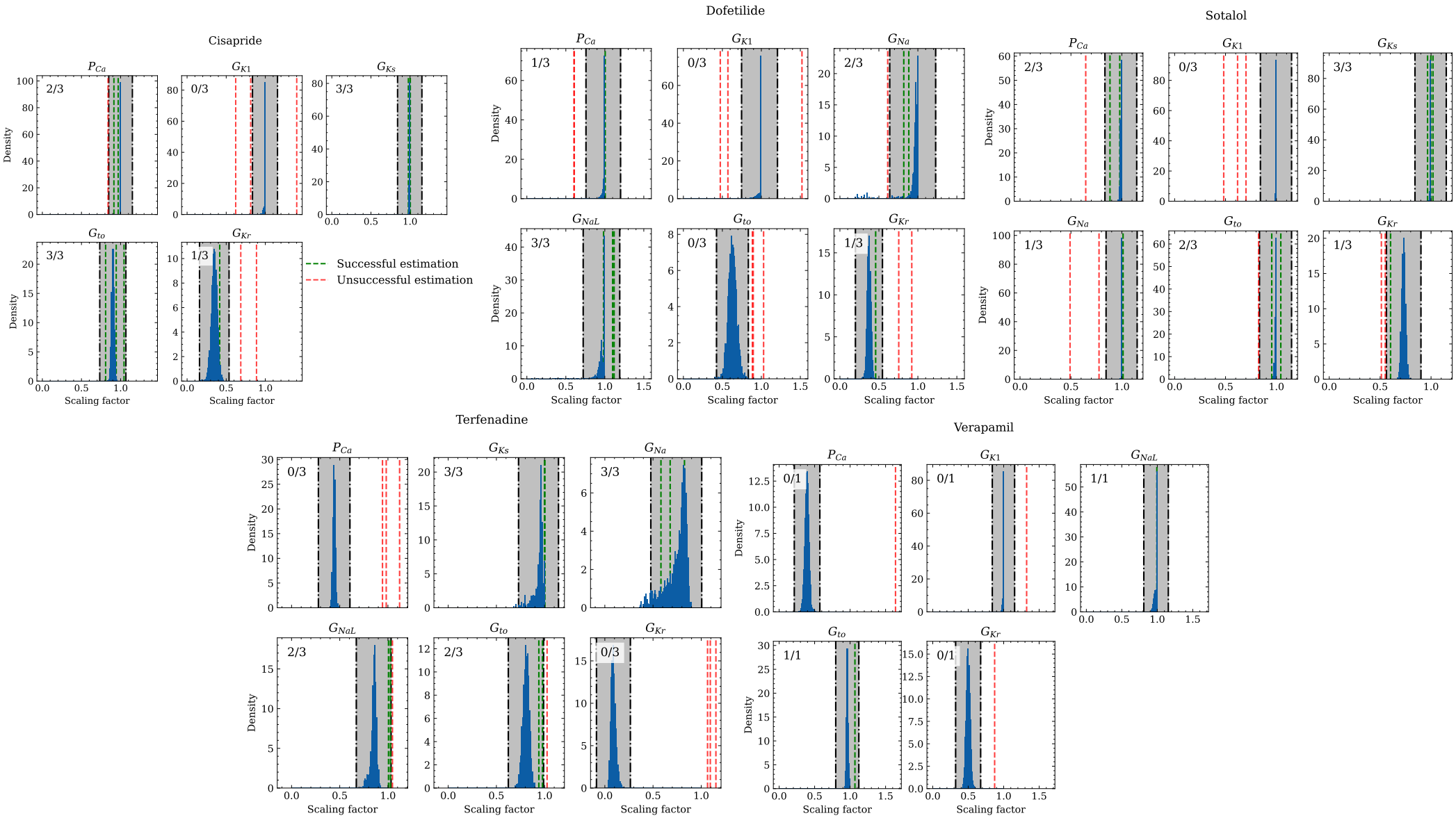

Figure 10: Analysis of solution accuracy of the inverse problem using experimental data. The histograms compare the estimated pharmacological parameters (dashed vertical lines) from data of multiple preparations with the CiPA distributions (blue; see Section 2.3.3). The black dash dotted vertical lines are the borders of the range (grey) that was used to determine if the estimation of the given pharmacological parameter was successful. The range is spanned up by $\mu \pm (0.15 + \sigma)$ of the CiPA distribution. Successfully estimated parameters are shown as green lines and unsuccessfully estimated parameters are shown as red lines. The number in the upper left corner indicates for how many preparations the parameters could be successfully estimated in relation to the total number of preparations for the given drug.

[Britton et al., 2013] Britton, O. J., Bueno-Orovio, A., Ammel, K. V., Lu, H. R., Towart, R., Gallacher, D. J., and Rodriguez, B. (2013). Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proceedings of the National Academy of Sciences*, 110(23).

[Britton et al., 2017] Britton, O. J., Bueno-Orovio, A., Virág, L., Varró, A., and Rodriguez, B. (2017). The electrogenic na^+/k^+ pump is a key determinant of repolarization abnormality susceptibility in human ventricular cardiomyocytes: A population-based simulation study. *Frontiers in Physiology*, 8.

[Chang et al., 2015] Chang, E. T. Y., Strong, M., and Clayton, R. H. (2015). Bayesian sensitivity analysis of a cardiac cell model using a gaussian process emulator. *PLOS ONE*, 10(6):e0130252.

[Chang et al., 2017] Chang, K. C., Dutta, S., Mirams, G. R., Beattie, K. A., Sheng, J., Tran, P. N., Wu, M., Wu, W. W., Colatsky, T., Strauss, D. G., and Li, Z. (2017). Uncertainty quantification reveals the importance of data variability and experimental design considerations for in silico proarrhythmia risk assessment. *Frontiers in Physiology*, 8.

[Clevert et al., 2016] Clevert, D.-A., Unterthiner, T., and Hochreiter, S. (2016). Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). arXiv:1511.07289 [cs].

[Colatsky et al., 2016] Colatsky, T., Fermini, B., Gintant, G., Pierson, J. B., Sager, P., Sekino, Y., Strauss, D. G., and Stockbridge, N. (2016). The comprehensive in vitro proarrhythmia assay (CiPA) initiative — update on progress. *Journal of Pharmacological and Toxicological Methods*, 81:15–20.

[Corrado et al., 2020] Corrado, C., Avezzù, A., Lee, A. W. C., Costa, C. M., Roney, C. H., Strocchi, M., Bishop, M., and Niederer, S. A. (2020). Using cardiac ionic cell models to interpret clinical data. *WIREs Mechanisms of Disease*, 13(3).

[Coveney and Clayton, 2018] Coveney, S. and Clayton, R. H. (2018). Fitting two human atrial cell models to experimental data using bayesian history matching. *Progress in Biophysics and Molecular Biology*, 139:43–58.

[Coveney et al., 2021] Coveney, S., Corrado, C., Oakley, J. E., Wilkinson, R. D., Niederer, S. A., and Clayton, R. H. (2021). Bayesian calibration of electrophysiology models using restitution curve emulators. *Frontiers in Physiology*, 12.

[Crumb et al., 2016] Crumb, W. J., Vicente, J., Johannesen, L., and Strauss, D. G. (2016). An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel. *Journal of Pharmacological and Toxicological Methods*, 81:251–262.

[Dutta et al., 2017] Dutta, S., Chang, K. C., Beattie, K. A., Sheng, J., Tran, P. N., Wu, W. W., Wu, M., Strauss, D. G., Colatsky, T., and Li, Z. (2017). Optimization of an in silico cardiac cell model for proarrhythmia risk assessment. *Frontiers in Physiology*, 8.

[Ghosh et al., 2018] Ghosh, S., Gavaghan, D. J., and Mirams, G. R. (2018). Gaussian process emulation for discontinuous response surfaces with applications for cardiac electrophysiology models.

[Grandi et al., 2010] Grandi, E., Pasqualini, F. S., and Bers, D. M. (2010). A novel computational model of the human ventricular action potential and ca transient. *Journal of Molecular and Cellular Cardiology*, 48(1):112–121.

[Groenendaal et al., 2015] Groenendaal, W., Ortega, F. A., Kherlopian, A. R., Zygmunt, A. C., Krogh-Madsen, T., and Christini, D. J. (2015). Cell-specific cardiac electrophysiology models. *PLOS Computational Biology*, 11(4):e1004242.

[Guillaume et al., 2019] Guillaume, J. H., Jakeman, J. D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M. C., Jakeman, A. J., Keesman, K. J., Razavi, S., and Stigter, J. D. (2019). Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modelling & Software*, 119:418–432.

[Hall et al., 2021] Hall, C., Gehmlich, K., Denning, C., and Pavlovic, D. (2021). Complex relationship between cardiac fibroblasts and cardiomyocytes in health and disease. *Journal of the American Heart Association*, 10(5).

[Herman and Usher, 2017] Herman, J. and Usher, W. (2017). SALib: An open-source python library for sensitivity analysis. *J. Open Source Softw.*, 2(9):97.

[Homma and Saltelli, 1996] Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Saf.*, 52(1):1–17.

[Hornik et al., 1989] Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366.

[Jæger et al., 2021] Jæger, K. H., Edwards, A. G., Giles, W. R., and Tveito, A. (2021). A computational method for identifying an optimal combination of existing drugs to repair the action potentials of SQT1 ventricular myocytes. *PLOS Computational Biology*, 17(8):e1009233.

[Jæger et al., 2019] Jæger, K. H., Wall, S., and Tveito, A. (2019). Detecting undetectables: Can conductances of action potential models be changed without appreciable change in the transmembrane potential? *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(7):073102.

[Jagtap et al., 2020] Jagtap, A. D., Kawaguchi, K., and Karniadakis, G. E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404:109136.

[Jeong et al., 2022] Jeong, D. U., Yoo, Y., Marcellinus, A., Kim, K.-S., and Lim, K. M. (2022). Proarrhythmic risk assessment of drugs by dvm /dt shapes using the convolutional neural network. *CPT: Pharmacometrics & Systems Pharmacology*, 11(5):653–664.

[Jeong et al., 2023] Jeong, D. U., Yoo, Y., Marcellinus, A., and Lim, K. M. (2023). Application of convolutional neural networks using action potential shape for in-silico proarrhythmic risk assessment. *Biomedicines*, 11(2):406.

[Johnstone et al., 2016] Johnstone, R. H., Chang, E. T., Bardenet, R., de Boer, T. P., Gavaghan, D. J., Pathmanathan, P., Clayton, R. H., and Mirams, G. R. (2016). Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models? *Journal of Molecular and Cellular Cardiology*, 96:49–62.

[Kingma and Ba, 2017] Kingma, D. P. and Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*. arXiv: 1412.6980.

[Kohl and Gourdie, 2014] Kohl, P. and Gourdie, R. G. (2014). Fibroblast–myocyte electrotonic coupling: Does it occur in native cardiac tissue? *Journal of Molecular and Cellular Cardiology*, 70:37–46.

[Krogh-Madsen et al., 2016] Krogh-Madsen, T., Sobie, E. A., and Christini, D. J. (2016). Improving cardiomyocyte model fidelity and utility via dynamic electrophysiology protocols and optimization algorithms. *The Journal of Physiology*, 594(9):2525–2536.

[Lei and Mirams, 2021] Lei, C. L. and Mirams, G. R. (2021). Neural Network Differential Equations For Ion Channel Modelling. *Frontiers in Physiology*, 12.

[Li et al., 2017] Li, Z., Dutta, S., Sheng, J., Tran, P. N., Wu, W., Chang, K., Mdluli, T., Strauss, D. G., and Colatsky, T. (2017). Improving the in silico assessment of proarrhythmia risk by combining hERG (human ether-à-go-go-related gene) channel–drug binding kinetics and multichannel pharmacology. *Circulation: Arrhythmia and Electrophysiology*, 10(2).

[Li et al., 2019] Li, Z., Ridder, B. J., Han, X., Wu, W. W., Sheng, J., Tran, P. N., Wu, M., Randolph, A., Johnstone, R. H., Mirams, G. R., Kuryshev, Y., Kramer, J., Wu, C., Crumb, W. J., and Strauss, D. G. (2019). Assessment of an in silico mechanistic model for proarrhythmia risk prediction under the CiPA initiative. *Clinical Pharmacology & Therapeutics*, 105(2):466–475.

[Mayer et al., 2017] Mayer, A., Bittihn, P., and Luther, S. (2017). Complex restitution behavior and reentry in a cardiac tissue model for neonatal mice. *Physiological Reports*, 5(19):e13449.

[Mirams et al., 2011] Mirams, G. R., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B. M., McMahon, N. C., Gavaghan, D. J., and Noble, D. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovascular Research*, 91(1):53–61.

[Mirams et al., 2014] Mirams, G. R., Davies, M. R., Brough, S. J., Bridgland-Taylor, M. H., Cui, Y., Gavaghan, D. J., and Abi-Gerges, N. (2014). Prediction of thorough QT study results using action potential simulations based on ion channel screens. *Journal of Pharmacological and Toxicological Methods*, 70(3):246–254.

[Mirams et al., 2016] Mirams, G. R., Pathmanathan, P., Gray, R. A., Challenor, P., and Clayton, R. H. (2016). Uncertainty and variability in computational and mathematical models of cardiac physiology. *The Journal of Physiology*, 594(23):6833–6847.

[Mohammadi et al., 2019] Mohammadi, H., Challenor, P., and Goodfellow, M. (2019). Emulating dynamic non-linear simulators using gaussian processes. *Computational Statistics & Data Analysis*, 139:178–196.

[Morotti et al., 2021] Morotti, S., Liu, C., Hegyi, B., Ni, H., Iseppe, A. F., Wang, L., Pritoni, M., Ripplinger, C. M., Bers, D. M., Edwards, A. G., and Grandi, E. (2021). Quantitative cross-species translators of cardiac myocyte electrophysiology: Model training, experimental validation, and applications. *Science Advances*, 7(47).

[Muszkiewicz et al., 2016] Muszkiewicz, A., Britton, O. J., Gemmell, P., Passini, E., Sánchez, C., Zhou, X., Carusi, A., Quinn, T. A., Burrage, K., Bueno-Orovio, A., and Rodriguez, B. (2016). Variability in cardiac electrophysiology: Using experimentally-calibrated populations of models to move beyond the single virtual physiological human paradigm. *Progress in Biophysics and Molecular Biology*, 120(1-3):115–127.

[Oakley and O'Hagan, 2004] Oakley, J. E. and O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: a bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):751–769.

[O'Hara et al., 2011] O'Hara, T., Virág, L., Varró, A., and Rudy, Y. (2011). Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLoS Computational Biology*, 7(5):e1002061.

[Orvos et al., 2019] Orvos, P., Kohajda, Z., Szlovák, J., Gazdag, P., Árpádffy-Lovas, T., Tóth, D., Geramipour, A., Tálosi, L., Jost, N., Varró, A., and Virág, L. (2019). Evaluation of possible proarrhythmic potency: Comparison of the effect of dofetilide, cisapride, sotalol, terfenadine, and verapamil on hERG and native i_{Kr} currents and on cardiac action potential. *Toxicological Sciences*, 168(2):365–380.

[Passini et al., 2017] Passini, E., Britton, O. J., Lu, H. R., Rohrbacher, J., Hermans, A. N., Gallacher, D. J., Greig, R. J. H., Bueno-Orovio, A., and Rodriguez, B. (2017). Human in silico drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Frontiers in Physiology*, 8.

[Passini et al., 2021] Passini, E., Zhou, X., Trovato, C., Britton, O. J., Bueno-Orovio, A., and Rodriguez, B. (2021). The virtual assay software for human in silico drug trials to augment drug cardiac testing. *Journal of Computational Science*, 52:101202.

[Plank et al., 2021] Plank, G., Loewe, A., Neic, A., Augustin, C., Huang, Y.-L., Gsell, M. A., Karabelas, E., Nothstein, M., Prassl, A. J., Sánchez, J., Seemann, G., and Vigmond, E. J. (2021). The openCARP simulation environment for cardiac electrophysiology. *Computer Methods and Programs in Biomedicine*, 208:106223.

[Rasmussen and Williams, 2019] Rasmussen, C. E. and Williams, C. K. I. (2019). *Gaussian processes for machine learning*. Adaptive Computation and Machine Learning Series. MIT Press, London, England.

[Rush and Larsen, 1978] Rush, S. and Larsen, H. (1978). A practical algorithm for solving dynamic membrane equations. *IEEE Transactions on Biomedical Engineering*, BME-25(4):389–392.

[Sager et al., 2014] Sager, P. T., Gintant, G., Turner, J. R., Pettit, S., and Stockbridge, N. (2014). Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the cardiac safety research consortium. *American Heart Journal*, 167(3):292–300.

[Saltelli, 2002] Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297.

[Sarkar and Sobie, 2010] Sarkar, A. X. and Sobie, E. A. (2010). Regression analysis for constraining free parameters in electrophysiological models of cardiac cells. *PLoS Computational Biology*, 6(9):e1000914.

[Sher et al., 2022] Sher, A., Niederer, S. A., Mirams, G. R., Kirpichnikova, A., Allen, R., Pathmanathan, P., Gavaghan, D. J., van der Graaf, P. H., and Noble, D. (2022). A quantitative systems pharmacology perspective on the importance of parameter identifiability. *Bulletin of Mathematical Biology*, 84(3).

[Shim et al., 2023] Shim, J. V., Xiong, Y., Dhanan, P., Dariolli, R., Azeloglu, E. U., Hu, B., Jayaraman, G., Schaniel, C., Birtwistle, M. R., Iyengar, R., Dubois, N. C., and Sobie, E. A. (2023). Predicting individual-specific cardiotoxicity responses induced by tyrosine kinase inhibitors. *Frontiers in Pharmacology*, 14.

[Sobol′, 2001] Sobol′, I. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280.

[Strauss et al., 2021] Strauss, D. G., Wu, W. W., Li, Z., Koerner, J., and Garnett, C. (2021). Translational models and tools to reduce clinical trials and improve regulatory decision making for QTc and proarrhythmia risk (ICH e14/s7b updates). *Clinical Pharmacology & Therapeutics*, 109(2):319–333.

[Syed et al., 2005] Syed, Z., Vigmond, E., Nattel, S., and Leon, L. J. (2005). Atrial cell action potential parameter fitting using genetic algorithms. *Medical and Biological Engineering and Computing*, 43(5):561–571.

[Tomek et al., 2019] Tomek, J., Bueno-Orovio, A., Passini, E., Zhou, X., Minchole, A., Britton, O., Bartolucci, C., Severi, S., Shrier, A., Virag, L., Varro, A., and Rodriguez, B. (2019). Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block. *eLife*, 8.

[Tomek et al., 2020] Tomek, J., Bueno-Orovio, A., and Rodriguez, B. (2020). Tor-ord-dyncl: an update of the tor-ord model of human ventricular cardiomyocyte with dynamic intracellular chloride. *bioRxiv*.

[Tveito et al., 2018] Tveito, A., Jæger, K. H., Huebsch, N., Charrez, B., Edwards, A. G., Wall, S., and Healy, K. E. (2018). Inversion and computational maturation of drug response using human stem cell derived cardiomyocytes in microphysiological systems. *Scientific Reports*, 8(1).

[Tveito et al., 2020] Tveito, A., Jæger, K. H., Maleckar, M. M., Giles, W. R., and Wall, S. (2020). Computational translation of drug effects from animal experiments to human ventricular myocytes. *Scientific Reports*, 10(1).

[Vicente et al., 2017] Vicente, J., Zusterzeel, R., Johannesen, L., Mason, J., Sager, P., Patel, V., Matta, M. K., Li, Z., Liu, J., Garnett, C., Stockbridge, N., Zineh, I., and Strauss, D. G. (2017). Mechanistic model-informed proarrhythmic risk assessment of drugs: Review of the "CiPA" initiative and design of a prospective clinical validation study. *Clinical Pharmacology & Therapeutics*, 103(1):54–66.

[Vigmond et al., 2008] Vigmond, E., dos Santos, R. W., Prassl, A., Deo, M., and Plank, G. (2008). Solvers for the cardiac bidomain equations. *Progress in Biophysics and Molecular Biology*, 96(1-3):3–18.

[Weiss et al., 2010] Weiss, J. N., Garfinkel, A., Karagueuzian, H. S., Chen, P.-S., and Qu, Z. (2010). Early afterdepolarizations and cardiac arrhythmias. *Heart Rhythm*, 7(12):1891–1899.

[Whittaker et al., 2020] Whittaker, D. G., Clerx, M., Lei, C. L., Christini, D. J., and Mirams, G. R. (2020). Calibration of ionic and cellular cardiac electrophysiology models. *WIREs Systems Biology and Medicine*, 12(4).

[Zaniboni et al., 2010] Zaniboni, M., Riva, I., Cacciani, F., and Groppi, M. (2010). How different two almost identical action potentials can be: A model study on cardiac repolarization. *Mathematical Biosciences*, 228(1):56–70.

6 Additional information

6.1 Funding

National Research, Development and Innovation Office Project, No. 142738
Eötvös Loránd Research Network
Norbert Jost
András Varró
German Research Foundation, Walter Benjamin Fellowship, No. 468256475
Alexander Jung
ERA-NET co-fund action no. 680969 funded by the Austrian Science Fund, grant I 4652-B
Christoph M. Augustin
Wellcome Senior Research Fellowship, No. 212203/Z/18/Z
Gary R. Mirams
The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

This research was funded in whole, or in part, by the Wellcome Trust [212203/Z/18/Z]. For the purpose of open access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

6.2 Author contributions

Thomas Grandits: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing
Christoph M. Augustin: Funding acquisition, Writing – review & editing
Gundolf Haase: Funding acquisition, Resources, Writing – review & editing
Norbert Jost: Funding acquisition, Resources, Writing – review & editing
Gary Mirams: Funding acquisition, Methodology, Writing – review & editing
Steven Niederer: Funding acquisition, Writing – review & editing
Gernot Plank: Funding acquisition, Resources, Writing – review & editing
László Virág: Resources, Writing – review & editing
András Varró: Funding acquisition, Resources, Writing – review & editing
Alexander Jung: Conceptualization, Data curation, Funding acquisition, Formal analysis, Investigation, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing

Appendices

A AP biomarkers and abnormalities

The AP biomarkers used in this study were selected such that the key characteristics of the depolarization and the repolarization phase can be quantified. They include RMP (resting transmembrane potential measured just before stimulation), $dVmMax$ (maximum transmembrane potential slope during the upstroke), $Peak$ (peak transmembrane potential at the end of the upstroke), APD_x (AP duration at $x \in \{30, 40, 50, 60, 70, 80, 90\}\%$ repolarization relative to the AP amplitude ($Peak - RMP$) measured from the instant of $dVmMax$), and Tri_{90-40} (triangulation defined as the difference between APD_{90} and APD_{40} [Britton et al., 2017]).
For the GSA (see Appendix B) and the creation of synthetic data (see Section 2.3.1), AP abnormalities were also considered which included depolarization abnormalities, repolarization abnormalities, and alternans. Depolarization abnormalities were defined as an upstroke peak below 0 mV and an AP that does not reach 0 mV before 100 ms after stimulation [Passini et al., 2017]. Repolarization abnormalities were defined as a transmembrane potential rate of rise of more than 0.01 $\frac{\text{mV}}{\text{ms}}$ from 150 ms after the upstroke peak onwards (representative of early afterdepolarizations) and as a transmembrane potential that does not fall below -40 mV [Passini et al., 2017] (representative of repolarization failure). Alternans were defined as APD_{90} difference of more than 5 ms between two consecutive APs [Morotti et al., 2021].

B Global sensitivity analysis

A variance-based Sobol' global sensitivity analysis (GSA) [Sobol′, 2001] was performed on the simulator to quantify the sensitivities of the maximum channel conductances (inputs) with respect to the AP biomarkers (outputs; see Appendix A). This informed the decision on which inputs to consider in the emulator. Furthermore, it was used for the interpretation of the solutions of the inverse problem since parameters that are insensitive with respect to the outputs indicate non-identifiability [Guillaume et al., 2019].

The maximum conductances used for building the model population in [Tomek et al., 2019] were considered and Saltelli's sampling scheme [Saltelli, 2002] was applied with $N = 1024$ to generate 20,480 input samples with values between 50% and 150% of the original values. Simulations were performed for each input sample and biomarker values (see Appendix A) derived from the last AP were used for the analysis. However, data were excluded if not all biomarkers could be determined or abnormalities (see Appendix A) were detected in the last two consecutive APs. First-order (S1) and total-effect (ST) Sobol' sensitivity indices were computed using the Saltelli method [Homma and Saltelli, 1996, Saltelli, 2002]. This requires outputs for each input sample and to take this into account, excluded outputs were assigned the mean values of included outputs. The `SALib-Sensitivity Analysis Library` [Herman and Usher, 2017] was used for the GSA.

The GSA could be performed on the data of all input samples as no data were excluded. S1 and ST were mostly very similar which indicates only little interactions among the maximum conductivities relative to the AP biomarkers (Figure 11). The only exception was TP. As was to be expected, the analysis underlines the predominant relative sensitivity of G_{Na} with respect to biomarkers of the depolarization phase, the predominant relative sensitivity of G_{NaL}, G_{Kr}, and G_{NCX} with respect to biomarkers of the repolarization phase and the predominant relative sensitivity of G_{K1}, P_{NaK} to the resting transmembrane potential. However, G_{Ks} has a negligible relative sensitivity to all biomarkers. This indicates almost non-identifiability.

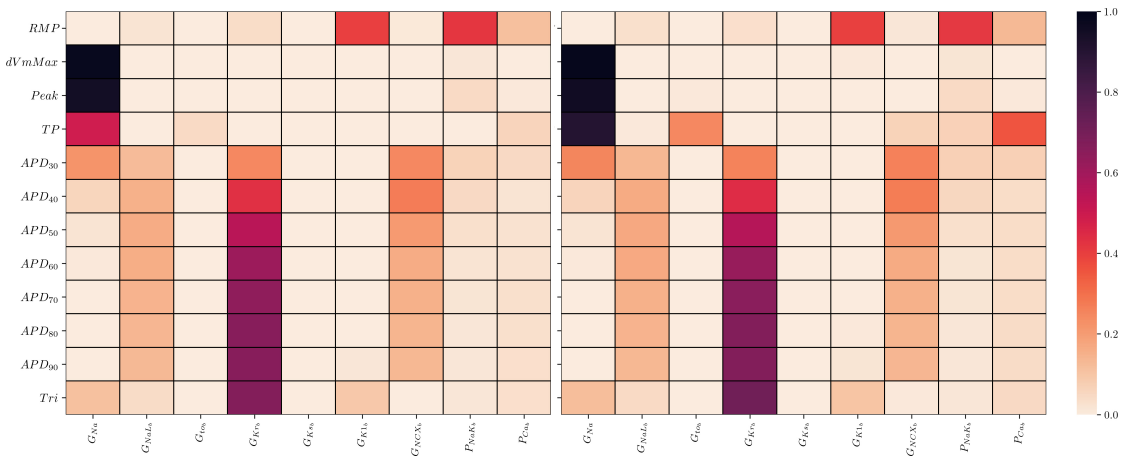

Figure 11: Global sensitivity analysis of the ToR-ORd simulator. Sobol' sensitivity indices are shown for each maximum conductance relative to each AP biomarker. Left: first-order (S1), right: total-effect (ST) Sobol' sensitivity coefficient.

C EAD classification

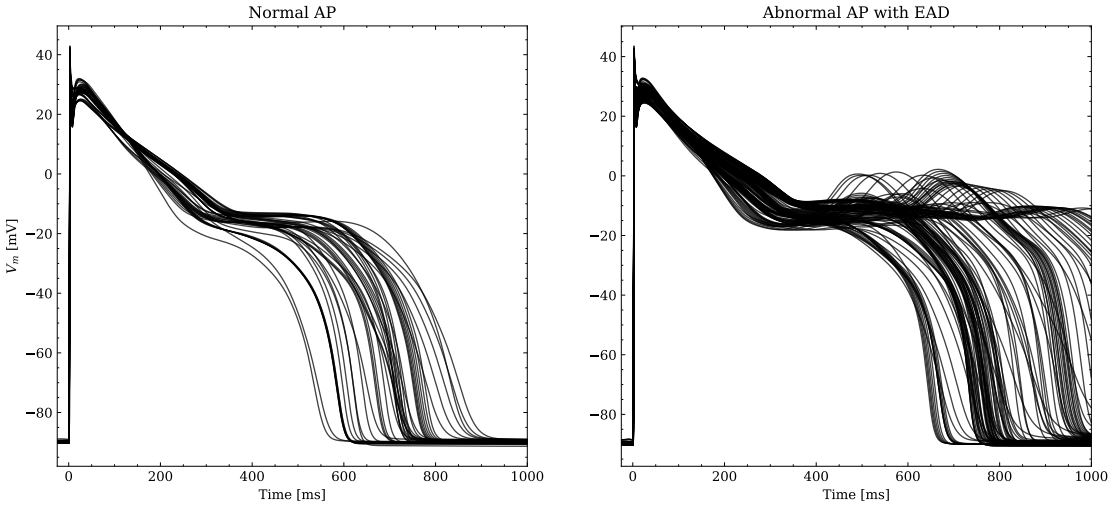

Figure 12: Emulated APs based on the pharmacological parameters of data set #2. See also Section 3.1.1 and Figure 7. Left: APs classified as normal, right: APs classified as abnormal exhibiting EADs. The classification criterion is outlined in Appendix A.