33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Regression methods for parameter sensitivity analysis: applications to cardiac arrhythmia mechanisms

Eric A. Sobie, Amrita X. Sarkar

Abstract—**Mathematical models are used extensively in studies of cardiac electrophysiology and arrhythmia mechanisms. Models can generate novel predictions, suggest experiments, and provide a quantitative understanding of underlying mechanisms. Limitations of present modeling approaches, however, include non-uniqueness of both parameters and the models themselves, and difficulties in accounting for experimental variability. We describe new approaches that can begin to address these limitations, and show how these can provide novel insight into mathematical models of cardiac myocytes.**

I. Introduction

Heart disease is one of the leading public health concerns in the United States. Among the causes of death in patients with heart disease, sudden cardiac death due to ventricular arrhythmia is the most common. Over the past several decades, researchers have made considerable progress in identifying the proteins that are altered in both congenital and acquired arrhythmia disorders. By and large, however, these basic science advances have not been translated into successful new therapies for arrhythmia treatment or prevention.

Cardiac electrical activity results from dynamic interactions between dozens of important proteins. The action potential (AP) results from the activity of Na^+ channels, Ca^{2+} channels, and several different types of K^+ channels, most of which exhibit complex time and voltage-dependences. The membrane also contains several electrogenic ion pumps and transporters that both regulate ion concentrations and generate ionic current that influences the AP. Alterations in virtually any of these proteins can influence the function of the others through changes in membrane voltage and intracellular [Ca^{2+}]. Moreover, cellular behaviors that can directly cause arrhythmias involve interactions between multiple participants. A systems-level mindset is therefore required to develop strategies to prevent these events.

As a result of this complexity, mathematical modeling has long been a common method for gaining insight into both normal and pathological electrophysiology [1]. Contemporary cell models typically consist of a coupled set of 15-60 ordinary differential equations (ODEs), each of which describes the temporal changes in a variable such as membrane voltage, an intracellular ionic concentration, or a dimensionless variable describing ion channel gating. A typical model also contains dozens of parameters with clear biological meaning such as the quantity of a particular type of ion channel, the speed with which that channel gates, or that channel's voltage dependence. These myocyte models successfully recapitulate measurable cellular variables such as AP shape, Ca^{2+} transients, and how these depend on pacing rate. A prominent example is the model of the guinea pig ventricular myocyte originally described by Luo & Rudy [2], then improved in several subsequent publications from this laboratory [3-5]. Models of this class have been used to provide insight into phenomena such as developmental differences in physiology [6], mechanisms of potentially arrhythmogenic early afterdepolarizations [7], and how mutations in ion channels lead to arrhythmogenic behavior in cells [8].

The vast majority of these studies, however, have adhered to the following paradigm. Changes in ionic current behavior resulting from a perturbation are hypothesized based on experimental data and the investigator's intuition. Simulations performed under these altered conditions are compared with model results obtained under control conditions, usually with the baseline, published parameters. If the effects of the perturbation on the simulation results are similar to the experimental observations, the changes are considered provisionally sufficient to explain the altered behavior. If the results do not match, additional changes to the model are sometimes considered.

It is important to note that investigations such as these suffer from several important limitations. (1) Only conditions that are explicitly considered can be understood. A number of other changes might hypothetically contribute to the observed behavior, but these possibilities remain unknown until circumstance dictates that they be addressed. Since contemporary models contain dozens of numerical parameters, the investigator's imagination may become the factor that limits which possibilities are investigated. (2) Although virtually all published models are validated against experimental data, the validation process is often incomplete and biased by whichever experimental studies are familiar to the investigators. Perhaps as a result of this, competing models meant to describe the same cell type may exhibit dramatically divergent behavior [9-12]. (3) Most

Manuscript received April 15, 2011.

This work was supported in part by U.S. National Institutes of Health grant GM071588 and a Grant-in-Aid from the American Heart Association, Heritage Affiliate (10GRNT4170020).

Both authors are with the Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY, USA. Correspondence to Dr. E.A. Sobie; phone 212-659-1706; email: eric.sobie@mssm.edu.

investigations consider the published model to represent a typical sample and ignore variability between individuals, which may influence the comparison between computations and experiments.

Here we describe our recent work that begins to address these limitations of the standard paradigm. We have developed novel methods to evaluate mathematical models of cardiac myocytes through systematic parameter sensitivity analyses. We illustrate how these methods can be used to generate novel and testable model predictions, to constrain model parameters based on a systematic comparison of model output with data, and to understand variability between individuals within a population. Methods such as these are likely are likely to become important for understanding the susceptibility of hearts to arrhythmias.

II. METHODS

The principle underlying our work is as follows. The traditional approach to both experimental and computational physiology is to examine effects one at a time. In contrast, we employ a relatively simple technique [10;13] based on parameter randomization, repeated simulations, and multivariable regression. This strategy examines changes in many model parameters all at once, thereby providing a systems-level view of myocyte function and overcoming many of the limitations of the traditional approach.

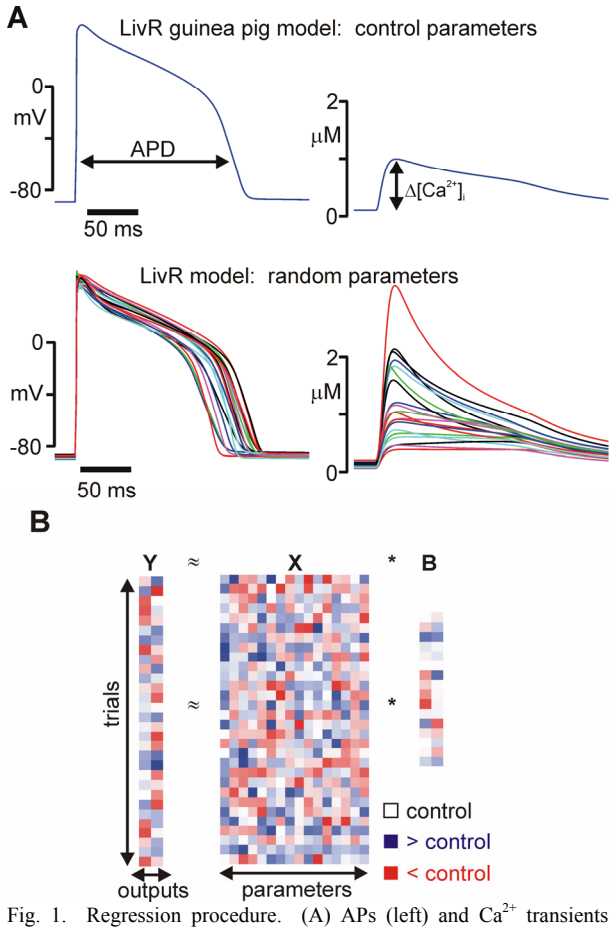

Fig. 1. Regression procedure. (A) APs (left) and Ca^{2+} transients (right) simulated with the LivR model. Standard parameters shown at top and 30 simulations with sets of random parameters shown at bottom. (B) Input and output matrices **X** and **Y** have the indicated structures. In this example, outputs are: 1) AP duration, 2) Ca^{2+} transient amplitude. The regression matrix **B** is derived such that **X*B** approximates **Y**.

Fig. 1 illustrates the method. With the published set of parameters, the LivR guinea pig model [4] produces the AP and Ca^{2+} transient shown in Fig. 1A, top. Randomizing the 16 ionic conductances in this model causes different output with each model variant (Fig. 1A, bottom). These randomly-varied parameters are placed in the input matrix **X**. (Note: we use the term 'conductances' generically, even though several parameters are not formally conductances. For instance, K_{SERCA}, the maximal turnover rate of the SR Ca^{2+} ATPase, is termed a conductance for convenience) With each simulation, physiologically-important metrics, such as AP duration (APD) and Ca^{2+} transient amplitude ($\Delta[Ca^{2+}]_i$) are computed and placed in the output matrix **Y** (Fig. 1B). Thus, each column of **X** corresponds to a different ionic conductance, and the two columns of **Y** (in this example) are APD and $\Delta[Ca^{2+}]_i$. Typically simulations are performed with hundreds of model variants. Multivariable regression computes a matrix **B** that minimizes the difference between **Y** and the matrix of predicted outputs $\hat{\mathbf{Y}} = \mathbf{X}*\mathbf{B}$. Our work [10;13] has shown that, despite many non-linearities in the underlying differential equations, linear approximations of relationships between parameters and outputs are usually surprisingly accurate. Moreover, although increasing the variability of the parameters in **X** decreases the accuracy of the regression model, this generally has minimal effects on the values of **B** [10].

III. RESULTS

A. *Importance of parameter sensitivities*

The values in the regression matrix **B** (Fig. 2) are parameter sensitivities – each value indicates how changing a particular parameter influences a specific model output. In contrast to the traditional approach of investigating parameters one-at-a-time, this procedure simultaneously provides information about many model parameters. The matrix **B** has two interpretations that are pertinent to the proposed work. First, each value in **B** represents a quantitative and testable model prediction, e.g. 70% inhibition of rapid delayed rectifier current I_{Kr} leads to a 20% increase in APD. Alternatively, since the conductances in each myocyte can be expected to be different, **B** can predict functional differences between individuals. Thus, if the row vector **x** represents how much the conductances deviate from control values in a given myocyte, then $\hat{\mathbf{y}} = \mathbf{x}*\mathbf{B}$ predicts how APD and $\Delta[Ca^{2+}]_i$ differ from the control values in that particular cell.

The parameter sensitivities (Fig. 2) show several

somewhat surprising predictions, even for a well-studied model system such as the guinea pig ventricular myocyte: (1) changes in conductances have greater effects (in percentage terms) on $\Delta[Ca^{2+}]_i$ than on APD (note different y-axis scales); (2) increases in background Ca^{2+} current (G_{CaB}), a depolarizing current, are predicted to shorten rather than lengthen APD; (3) changes in Na^+-K^+ pump activity (K_{NaK}) are predicted to have greater effects on $\Delta[Ca^{2+}]_i$ than changes in SERCA pump activity (K_{SERCA}). Most these predictions can be tested experimentally by applying drugs that selectively inhibit particular ion transport mechanisms and quantifying the change in the cellular response (i.e. APD or Ca^{2+} transient amplitude).

By examining model parameters all-at-a-time rather than one-at-a-time, this method generates a rich set of predictions that can guide experiments based on which experimental tests are most likely to generate interesting results. Moreover, a strong model should be able to recapitulate not just the baseline behavior of the myocyte, but also to successfully predict how the cell responds to a wide range of perturbations. By generating a large number of predictions all at once, this type of analysis provides a framework for a systematic comparison of model output with experimental data. If such analyses become a standard component of model development and testing, this can help overcome the limitation that model validation is often limited and biased.

B. Examining additional parameters and outputs

The parameters shown in Fig. 2 represent ionic-current maximal conductances and maximal rates of ion transport of pumps and transporters. In recent publications, we have extended this procedure to include additional parameters that

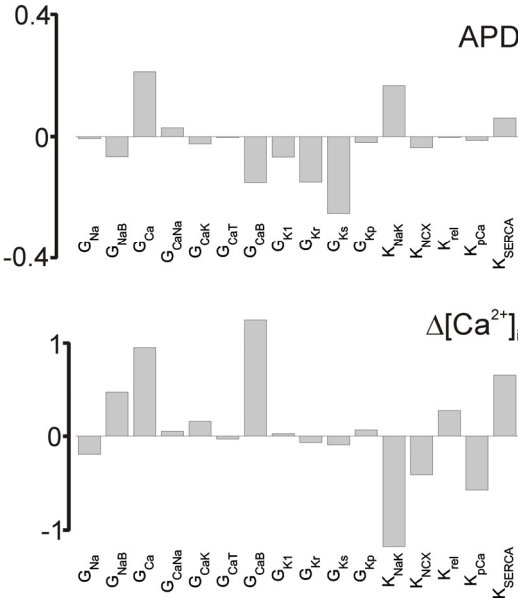

Fig. 2. Regression-derived parameter sensitivities indicate how much each conductance affects AP duration (top) or Ca^{2+} transient amplitude (bottom).

control either the kinetics of ion channel gating or the voltage-dependences of channel activation and inactivation [10]. This is relevant because it allows for a direct comparison between the predicted quantitative effects of a change in ion channel abundance versus a change in channel kinetics of voltage-dependence. Mutations, for instance, that may affect both abundance and channel properties such as kinetics can be interpreted in this framework.

Additionally, APD and $\Delta[Ca^{2+}]_i$ were chosen for this example because these commonly-measured variables have obvious physiological relevance. The method has successfully analyzed, however, more abstract outputs such as stimulation threshold, and outputs directly related to arrhythmias, such as the pacing rate at which APD alternans first appears [10;13].

A simple extension of this technique can be implemented to relax the assumption that changes in model parameters cause independent effects on outputs. Additional columns, consisting of the products of individual parameters, can be added to the matrix **X.** This approach can be used to explore potential nonlinear interactions between parameters.

C. Understanding the behavior of a population

The parameter sensitivity analysis method described above is useful for understanding how variability within a

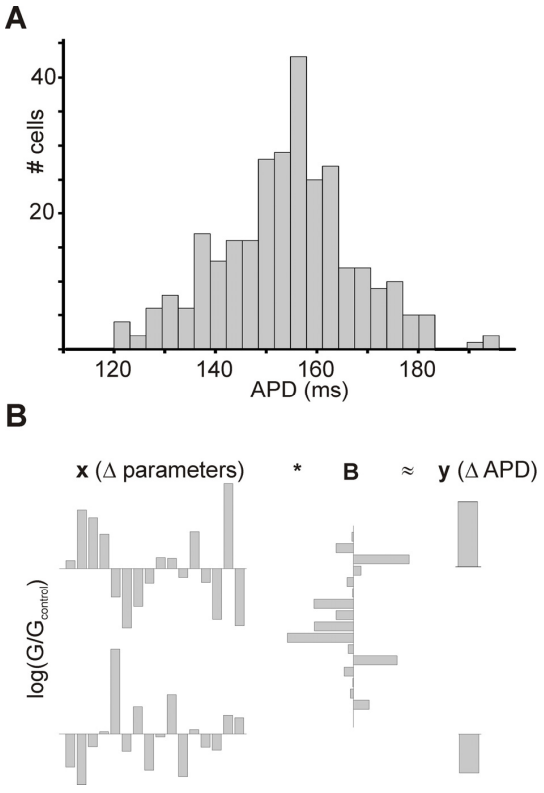

Fig. 3. Variability within a population of candidate models. (A) Distribution of APD in 296 LivR guinea pig myocytes with randomly varying parameters. (B) In two particular myocytes, the APD (expressed relative to control) can be predicted from the dot product of the change in parameters (expressed in logarithmic units relative to control, where increases are positive and decreases are negative) and the vector of parameter sensitivities. The top and bottom cells are predicted to have APDs that are greater than and less than control, respectively.

population in ion channel expression or function translates into variability in function. Remember that the initial step in the analysis is to run simulations with multiple candidate models, thereby generating a large set of pseudo-data. Fig. 3 illustrates how this can be exploited to understand differences between individuals. Fig. 3A shows the variability in APD within the sample of 300 candidate models, and Fig. 3B illustrates how the regression model can predict the behavior of individual myocytes within the population. The matrix formulation of the regression problem provides an extremely convenient framework for this purpose. For any particular myocyte, the change (relative to control) of a model output such as APD can be computed as the dot product of that cell's parameters and a vector of parameter sensitivities. This relationship follows from the equation $\hat{\mathbf{Y}} = \mathbf{X}*\mathbf{B}$. Fig. 3B illustrates this schematically for two myocytes selected from the population of 300.

D. Constraining model parameters.

The strong correlations observed between parameters and most outputs has the important consequence that the reverse procedure can be performed. That is, a set of experimentally-measurable metrics can be used to determine the model parameters required to produce those outputs. Mathematically, this procedure merely requires swapping the input and output matrices and performing the "reverse regression" such that a close approximation of the parameters $\mathbf{X}$ can be calculated as $\mathbf{X} \approx \mathbf{Y}\mathbf{B}_{reverse}$. In a recent publication [13] we demonstrated that this procedure can successfully constrain conductances in several mathematical models of cardiac myocytes. This work, along with other recent work from our group [14], represents a formalization of the intuitive principle that a model's parameters are more likely to be uniquely constrained when the model is asked to reproduce many, rather than just a few, experimental results.

IV. Conclusion

Methods to systematically evaluate mathematical models of heart cells are likely to become increasingly important for increasing the rigor of the model-experiment comparison and to generate new physiological insight in coming years. We have illustrated several possible applications of such techniques, including: (1) generation of novel predictions through all-at-a-time simulation tests; (2) treatments of heterogeneous populations of myocytes, and (3) methods for automatically constraining model parameter via a thorough comparison with data.

References

[1] D. Noble and Y. Rudy, "Models of cardiac ventricular action potentials: iterative interaction between experiment and simulation," *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, vol. 359, no. 1783, pp. 1127-1142, 2001.

[2] C. H. Luo and Y. Rudy, "A dynamic model of the cardiac ventricular action potential. I. Simulations of ionic currents and concentration changes," *Circ. Res.*, vol. 74, pp. 1071-1096, 1994.

[3] G. M. Faber, J. Silva, L. Livshitz, and Y. Rudy, "Kinetic properties of the cardiac L-type Ca^{2+} channel and its role in myocyte electrophysiology: a theoretical investigation," *Biophys. J.*, vol. 92, no. 5, pp. 1522-1543, Mar.2007.

[4] L. Livshitz and Y. Rudy, "Uniqueness and stability of action potential models during rest, pacing, and conduction using problem-solving environment," *Biophys. J.*, vol. 97, no. 5, pp. 1265-1276, Sept.2009.

[5] J. Zeng, K. R. Laurita, D. S. Rosenbaum, and Y. Rudy, "Two components of the delayed rectifier K+ current in ventricular myocytes of the guinea pig type. Theoretical formulation and their role in repolarization," *Circ. Res.*, vol. 77, no. 1, pp. 140-152, July1995.

[6] L. J. Wang and E. A. Sobie, "Mathematical model of the neonatal mouse ventricular action potential," *Am J Physiol Heart Circ Physiol*, vol. 294, no. 6, p. H2565-H2575, June2008.

[7] J. Zeng and Y. Rudy, "Early afterdepolarizations in cardiac myocytes: mechanism and rate dependence," *Biophys. J.*, vol. 68, no. 3, pp. 949-964, Mar.1995.

[8] C. E. Clancy and Y. Rudy, "Linking a genetic defect to its cellular phenotype in a cardiac arrhythmia," *Nature.*, vol. 400, no. 6744, pp. 566-569, Aug.1999.

[9] E. M. Cherry and F. H. Fenton, "A tale of two dogs: analyzing two models of canine ventricular electrophysiology," *Am J Physiol Heart Circ Physiol*, vol. 292, no. 1, p. H43-H55, Jan.2007.

[10] E. A. Sobie, "Parameter sensitivity analysis in electrophysiological models using multivariable regression," *Biophys. J*, vol. 96, no. 4, pp. 1264-1274, Feb.2009.

[11] K. H. Ten Tusscher, O. Bernus, R. Hren, and A. V. Panfilov, "Comparison of electrophysiological models for human ventricular cells and tissues," *Prog. Biophys. Mol. Biol.*, vol. 90, no. 1-3, pp. 326-345, Jan.2006.

[12] R. Wilders, "Computer modelling of the sinoatrial node," *Med. Biol. Eng Comput.*, vol. 45, no. 2, pp. 189-207, Feb.2007.

[13] A. X. Sarkar and E. A. Sobie, "Regression analysis for constraining free parameters in electrophysiological models of cardiac cells," *PLoS Comput. Biol.*, vol. 6, no. 9, p. e1000914, 2010.

[14] E. A. Sobie and H. R. Ramay, "Excitation-contraction coupling gain in ventricular myocytes: insights from a parsimonious model," *J Physiol*, vol. 587, no. Pt 6, pp. 1293-1299, Mar.2009.