Computer Aided Clinical Trials for Implantable Cardiac Devices

Kuk Jin Jang*, James Weimer†, Houssam Abbas *, Zhihao Jiang *,
Jackson Liang ‡, Sanjay Dixit ‡, Rahul Mangharam*

* Department of Electrical and Systems Engineering
University of Pennsylvania

† Department of Computer and Information Science
University of Pennsylvania

‡ Hospital of the University of Pennsylvania

***Abstract*—In this paper we aim to answer the question, "How can modeling and simulation of physiological systems be used to evaluate life-critical implantable medical devices?" Clinical trials for medical devices are becoming increasingly inefficient as they take several years to conduct, at very high cost and suffer from high rates of failure. For example, the Rhythm ID Goes Head-to-head Trial (RIGHT) sought to evaluate the performance of two arrhythmia discriminator algorithms for implantable cardioverter defibrillators, Vitality 2 vs. Medtronic, in terms of time-to-first inappropriate therapy, but concluded with results contrary to the initial hypothesis - after 5 years, 2,000+ patients and at considerable ethical and monetary cost. In this paper, we describe the design and performance of a computer-aided clinical trial (CACT) for Implantable Cardiac Devices where previous trial information, real patient data and closed-loop device models are effectively used to evaluate the trial with high confidence. We formulate the CACT in the context of RIGHT using a Bayesian statistical framework. We define a hierarchical model of the virtual cohort generated from a physiological model which captures the uncertainty in the parameters and allows for the systematic incorporation of information available at the design of the trial. With this formulation, the CACT estimates the inappropriate therapy rate of Vitality 2 compared to Medtronic as 33.22% vs 15.62% (p<0.001), which is comparable to the original trial. Finally, we relate the outcomes of the computer-aided clinical trial to the primary endpoint of RIGHT.**

I. Introduction

Medical device clinical trials (CTs) are time consuming and costly endeavors, where a late-phase trial can take over 4-6 years and cost a device manufacturer over $10-20 million [1]. Despite these costs, rigorously planned clinical trials obtain undesirable results for reasons such as assuming the incorrect effect direction or having inadequate power [2]. Such events motivate the need for alternative methods to test and evaluate medical cyber-physical systems.

Due to the increased complexity of medical devices, and the practical and ethical limitations in the design of medical device trials, CTs are becoming increasingly insufficient in evaluating the risk of new technologies. This leads to the question we aim to answer in this work, "Can simulated data be used to evaluate medical devices?"

The implantable cardioverter defibrillator (ICD) is an example of a medical device which diagnoses ventricular tachycardia (VT) and delivers therapy in the form of electrical shocks to terminate ventricular tachycardia and prevent sudden cardiac death. While ICDs have been show to reduce the mortality rate by up to 31% [3], they suffer from *inappropriate therapy* - delivery of unnecessary electric shocks due to misclassifying supraventricular tachycardias (SVTs) for VTs. Inappropriate therapy increases patient morbidity and stress, reduces their quality of life, and is linked to increased morbidity [4].

To demonstrate the computer-aided clinical trial (CACT) approach, we use the example of The Rhythm ID Goes Head-to-head Trial (RIGHT) - a 2187-patient CT, from 2005-2010. RIGHT sought to compare the VT/SVT arrhythmia discrimination algorithms used by two ICD models from Vitality II (V2) and Medtronic (MDT) with regards to the time-to-first inappropriate therapy. At the conclusion of the CT, the effect direction and size for the performance was *opposite* of what was hypothesized, with V2 ICDs having a 34% increase in the risk of inappropriate therapy compared to MDT ICDs.

Related Work. Regulatory institutions, such as the U.S. Food and Drug Administration (FDA) have previously recognized that computer-aided modeling and simulation have can possibly complement traditional CTs and act as various indications about the performance of a new treatment or device, such as the ICD [5][6]. Another example is the T1 Diabetes Mellitus Metabolic Simulator (T1DMS) of UVA/PADOVA[7]. The objective of the T1DMS model is to test the efficacy of new glucose control algorithms by simulating them on the virtual cohort which has a fixed virtual cohort with 300 patients. The T1DMS models glucose kinetics in hypoglycemia, and has been accepted by the FDA as a substitute for animal trials.

Extensive work has been presented regarding the incorporation of prior information into the design of a prospective trial in the form of historical trials [8][9] or by using stochastic engineering models as priors in a Bayesian clinical trial setting [10].

In this paper we describe a high-confidence CACT for ICDs. Fig. 1 depicts the overall structure of a CACT for medical devices. After defining the endpoint with respect to a CT, called the *target CT*, a synthetic *virtual cohort* of physiological signals input to the ICD, called electrograms (EGMs), are generated through simulation of instances of a heart model. The structure and parameters of this physiological model are derived from real patient data from the Ann Arbor Electrogram

Libraries and existing information about physiology. The synthetic EGMs are applied to a device model/algorithm and the outcomes evaluated.

Despite the utility of the CACT, the conclusions drawn from the results were limited due to uncertainty in the generated EGM. The performance of the device depends heavily on the overall characteristics of various prognostic factors in the synthetic cohort, such as the distribution of the occurrence of various arrhythmia types and the ventricular cycle length (VCL), the peak-to-peak interval of the ventricular EGM.

Contributions. Addressing the shortcomings of the prior work[11], we demonstrate how using the prior information as that was available at the time of RIGHT vastly improve the CACT outcomes to estimates of the inappropriate therapy rate. Specifically, we make the following contributions:

1) We capture the uncertainty in the parameters of the physiological model with a Bayesian hierarchical model. We incorporate prior information regarding the ranges VCL (e.g SVT $426 \pm 57[ms]$) available before RIGHT.
2) From the formulation, we obtain estimates of the inappropriate therapy rate of V2 compared to MDT devices (33.22% vs 15.62%), comparable to the original trial.
3) Finally, from the estimate of inappropriate therapy, we obtain an approximation of the time-to-inapproapriate therapy resulting in survival curves comparable to RIGHT (CACT vs. RIGHT hazard ratio: 1.30 vs. 1.63)

In Sec. II, we describe the formulation of the CACT within a Bayesian statistical framework in the context of RIGHT. In Sec. III, we describe the hierarchical model for the virtual cohort and method of evaluation. We report the results and limitations of our methods in Sec. IV. Finally, we conclude with possible recommendations/ideas for future work.

II. Problem Formulation

In this section, we formulate the general problem considered in this work. The primary endpoint of RIGHT was the time to first inappropriate therapy. The time-to-first inappropriate therapy is difficult to measure in simulation due to the long-term variability in the occurrence of arrhythmia at both the individual and the population level. Thus, we define the endpoint as the inappropriate therapy rate, θ, as the endpoint for the CACT, which is measured with individual EGMs.

A fundamental aspect of a medical device CACT stems from the fact that medical devices react to sensed physiological signals, where most (if not all) physiological models are governed by physiological parameters. Thus, uncertainty in underlying physiological parameters manifests as uncertainty in physiological signals, which should be considered within the CACT framework.

We model the uncertainty in a CACT using a Bayesian hierarchical model relating the signals generated in a virtual cohort, X_j, to the physiological parameters or *settings* of the physiological model, η, denoted as

$$X_j \sim p_X(x_j \mid \eta). \tag{1}$$

for $j = 1...N_o$, where N_o is the size of the cohort.

We denote the uncertainty in the parameter η using a prior distribution, parameterized by condition-specific available information, γ_t, where t denotes the heart condition, such that

$$\eta \sim p_\eta(\eta \mid \gamma_t). \tag{2}$$

In RIGHT, the physiological signals are the EGMs and the condition-specific available information is the VCL for each type of heart condition, $t \in$\{normal sinus rhythm (NSR), SVT, VT\}. In this work, we aim to utilize the prior information to aid in the formulation of the CACT such that uncertainties in the outcome can be made explicit and relate the conclusions back to the target CT. Towards this goal, we make the following assumptions:

1) We assume that the physiological model in [11] is probabilistic and captures the variability in EGMs cycle length for each heart condition.
2) A group of settings will produce a type of EGMs for a heart condition and the distribution of these groups of settings are independent conditioned on the type. The distribution of the settings is parameterized according to the VCL of the rhythm.
3) For the device model, we assume that the performance of the device with regards to inappropriate therapy rate, θ, will be fixed and the same for all devices of the same model, $d \in$ {V2,MDT}.

III. CACT with Bayesian Hierarchical Models

In this section, we describe the hierarchical model which explicitly states the uncertainty in the physiological signal and incorporate condition-specific available information through this prior. With this model, evaluate the outcomes of the medical device and estimate the inappropriate therapy rate.

Virtual Cohort Generation with Prior Distribution for Model Parameters The uncertainty in the EGM stems from the uncertainty in η can modeled using a prior distribution which is parameterized by the VCL for a rhythm, γ_t. We define a prior distribution on γ_t which is conditioned on the information in the literature the VCL for various types of conditions, D.

$$\gamma_t \sim p_{\gamma_t}(\gamma_t \mid D) \tag{3}$$

In this context, D is a value of the parameters which determine the frequency of a particular type of heart condition and the mean and standard deviation of the ranges of VCL for a condition. For example, the ranges of the VCL for VT in a clinical trial conducted prior to RIGHT was $257.3 \pm 41.2[ms]$. We define the complete prior distribution for EGM as:

$$X_j \sim p_X(x_j \mid \eta) p_\eta(\eta \mid D) \tag{4}$$

where, for notation, we have defined η to include γ_t and the prior as $p_\eta(\eta \mid D)$. To account for the uncertainty in the parameters we integrate:

$$p_X(x \mid D) = \int p_X(x \mid \eta) p_\eta(\eta \mid D) d\eta \tag{5}$$

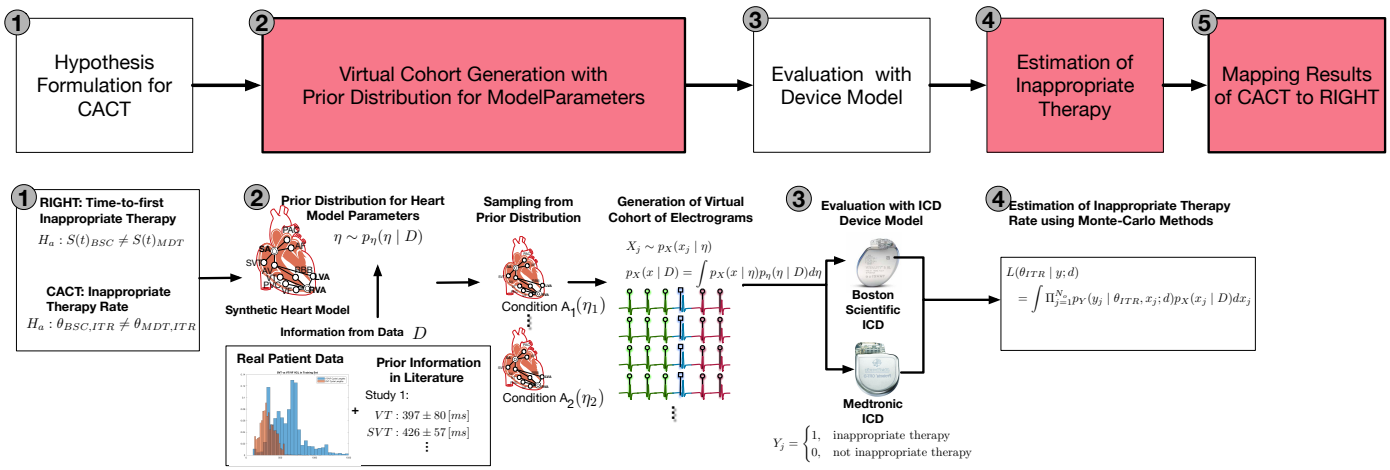

Figure 1: Overall structure of a CACT for ICDs. Contributions of this work are highlighted in red. In a CACT, real patient data and prior information is used to generate a synthetic cohort of EGM. The performance of the device model is evaluated based on the outcome of applying the cohort.

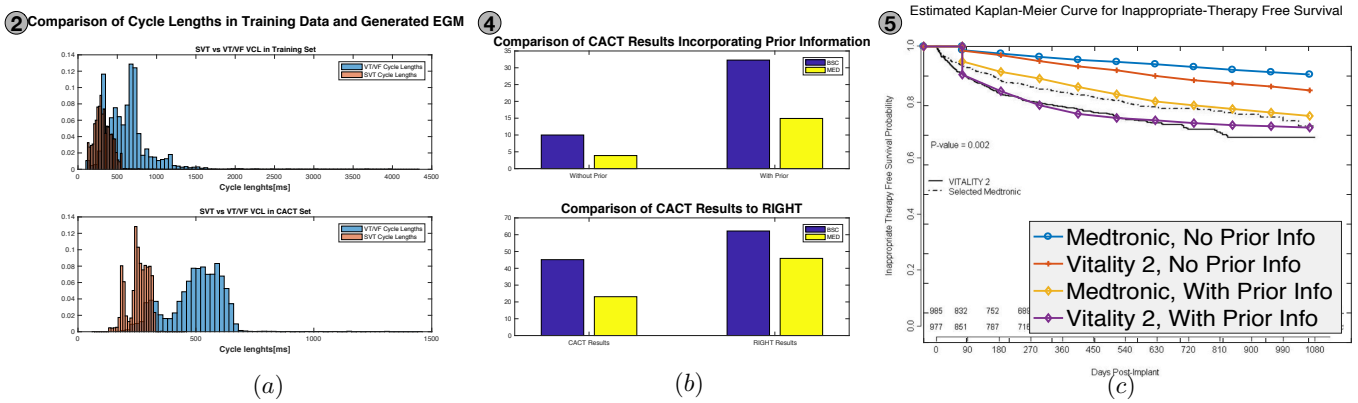

Figure 2: Main results: (a) Distribution of VCL for generated virtual cohort (b) Comparison of Inappropriate Therapy Rates (top) Comparison of CACT inappropriate therapy rate estimate with and without prior information. Uniform distribution of arrhythmia type assumed (bottom) Comparison of CACT estimate of inappropriate therapy rate to results from RIGHT. (c) Estimated RIGHT Survival Curves

Device Model - ICD Discrimination Algorithms

The output of the device model on the jth EGM in the cohort is modeled as independent Bernoulli random variable Y_j with parameters θ_d ($d \in \{VT, MDT\}$), and X_j. Y_j takes the following values:

$$Y_j = \begin{cases} T, & \text{output is inappropriate therapy} \\ F, & \text{output is not inappropriate therapy} \end{cases} \quad (6)$$

Estimate of Inappropriate Therapy Rate

From the responses of the device model, we can define the likelihood function for a device as:

$$L(\theta_d \mid y_o) = \Pi_{j=1}^{N_o} p_Y(y_j \mid \theta_d, x_j) p_{X_j}(x_j \mid D) \quad (7)$$

where x_o, y_o is a shorthand notation for the N_o EGMs and device outputs. Similar to (5), the variability in the generated cohort can be integrated out:

$$L(\theta_d \mid y) = \int \Pi_{j=1}^{N_o} p_Y(y_j \mid \theta_d, x_j) p_X(x_j \mid D) dx_j \quad (8)$$

There is no a closed-form solution for this expression, however we use Monte-Carlo methods to estimate θ [12].

IV. Results and Discussion

Effect of Incorporating Prior Data on VCL Distribution

Fig. 2(a) depicts the VCL of SVT and VT in both the real patient training data set and within the generated virtual cohort. Information available in [13] was incorporated according to the (5). The mean and standard deviation for one instance of the generated virtual cohort was $485.9 \pm 118[ms]$ for SVT and $258.0 \pm 43[ms]$ for VT. This is in comparison to the mean and $625.55 \pm 281.27[ms]$ for SVT and $322.10 \pm 137.11[ms]$ for VT within the real patient training data set. This confirms that the virtual cohort was generated reflected information that would have been available at the design of RIGHT.

Estimates of Inappropriate Therapy Rate We generated a virtual cohort with prior information incorporated, with 11,400 total EGMs (19 different conditions, 600 heart model instances simulated for 50 seconds) and obtained the corresponding responses. The generation procedure was repeated for a total of

100 iterations to obtain an overall estimate of the inappropriate therapy rate. By only assuming a uniform distribution on the occurrence of arrhythmia within a cohort, analogous to what would be assumed at the design stage of RIGHT, we obtained results that the V2 discrimination algorithm had a higher rate of inappropriate therapy than the MDT discrimination algorithm (33.22% vs 15.62% with p-value <0.001). Without utilizing the prior information, as in [11], the results are statistically significant, but the difference in effect size is not as pronounced (9.99% vs 3.88% with p-value <0.001). This would lead to greater uncertainty from the results. This comparison in shown in Fig. 2 (b,top).

In order to further validate the CACT outcomes, the distribution heart conditions from the results of RIGHT [1] was utilized to estimate the inappropriate therapy rate, retrospectively. With this additional information, effect sizes of 45.6% vs 23.11% for V2 vs. MDT. Thus, further reducing the gap between the results of the CACT and the results of RIGHT which were 62.2% vs 45.9% (Fig. 2 (b,bottom)).

Mapping CACT Results to RIGHT Assuming that the distribution of heart conditions remain constant, the time-to-first inappropriate therapy, the original endpoint of RIGHT, can be estimated using a geometric distribution, whose parameter is the inappropriate therapy rate, θ. We utilize the study in [14], to obtain information about the distribution of heart conditions and the rate of occurrence within a population, for example, for SVT: From the cohort of 1514 enrolled patients, 428 had 2596 non-ventricular SVT episodes, assuming a constant rate of occurrence, the average SVT per patient is:

$$\text{SVT per patient} = 2596/428 = 6.0654$$

$$\text{Interval of SVT events} = 3\,years * 365/\text{SVT per patient} = 180.5 \text{ [days]}$$

Sampling from the geometric distribution with parameter set to θ for each device, we plot these according to time using the values above and obtain the survival curves in Fig. 2 (c). The shape of the estimated curve is comparable to the actual survival curve of RIGHT. The estimated curve without the information incorporated exhibits very apparent differences in the shape.

After fitting a Cox proportional hazard curve to the estimates, we obtained a hazard ratio of 1.29677 (P<0.001) This was in comparison to a hazard ratio of 1.63 (P<0.001) obtained in RIGHT. The difference in the hazard ratio obtained with the prior information incorporated is significantly less compared to the discrepancy in hazard ratio without the prior information, which was 2.196 (P<0.001).

Limitations

One reason for discrepancy in the final effect sizes, is that the VCL for the various arrhythmia was not reported in [1]. If the VCL had been matched, as well as the distribution of arrhythmia, the estimate of the inappropriate therapy rate may have been more accurate.

Moreover, strong assumptions of the model, such as independence between model settings conditioned on the heart condition type, will not necessarily hold true in reality.

V. Conclusion and Future Work

In this work we formulated the CACT for ICDs within the context of a retrospective CT (RIGHT) and defined a hierarchical model which allows for incorporation of prior information available at the design of RIGHT. We obtained estimates of the inappropriate therapy rate comparable to RIGHT. Additionally, we approximate the the original endpoint of RIGHT, the time-to-inappropriate therapy, using the estimated inappropriate therapy rate.

Future work includes the usage of more sophisticated probabilistic models which account for other sources of uncertainty, as well as generalization to other types of medical devices.

References

[1] M. R. Gold, S. Ahmad, K. Browne, K. C. Berg, L. Thackeray, and R. D. Berger, "Prospective comparison of discrimination algorithms to prevent inappropriate ICD therapy: Primary results of the Rhythm ID Going Head to Head Trial ," *Heart Rhythm*, vol. 9, no. 3, pp. 370 – 377, 2012.

[2] M. D, D. CS, and W. GA, "Statistical power, sample size, and their reporting in randomized controlled trials," *JAMA*, vol. 272, no. 2, pp. 122–124, 1994.

[3] A. J. Moss *et al.*, "Reduction in inappropriate therapy and mortality through icd programming," *New England Journal of Medicine*, vol. 367, no. 24, pp. 2275–2283, 2012.

[4] M. Rosenqvist, T. Beyer, M. Block, K. Dulk, J. Minten, and F. Lindemans, "Adverse Events with Transvenous Implantable Cardioverter-Defibrillators: A Prospective Multi-center Study," *Circulation*, 1998.

[5] O. Faris and J. Shuren, "An fda viewpoint on unique considerations for medical-device clinical trials," *New England Journal of Medicine*, vol. 376, no. 14, pp. 1350–1357, 2017.

[6] Food and D. Administration, "Summary of safety and effectiveness data for pma p090013," FDA, Silver Spring, MD, Tech. Rep., 2011.

[7] C. Toffanin, M. Messori, F. Di Palma, G. De Nicolao, C. Cobelli, L. Magni, "Artificial Pancreas: Model Predictive Control Design from Clinical Experience," *J Diabetes Sci Technol*, vol. 7, pp. 1470–1483, 2014.

[8] J. G. Ibrahim, M.-H. Chen, Y. Gwon, and F. Chen, "The power prior: theory and applications," *Statistics in medicine*, vol. 34, no. 28, pp. 3724–3749, 2015.

[9] K. Viele, S. Berry, B. Neuenschwander, B. Amzal, F. Chen, N. Enas, B. Hobbs, J. G. Ibrahim, N. Kinnersley, S. Lindborg *et al.*, "Use of historical control data for assessing treatment effects in clinical trials," *Pharmaceutical statistics*, vol. 13, no. 1, pp. 41–54, 2014.

[10] T. Haddad, A. Himes, and M. Campbell, "Fracture prediction of cardiac lead medical devices using bayesian networks," *Reliability engineering & system safety*, vol. 123, pp. 145–157, 2014.

[11] Z. Jiang, H. Abbas, K. J. Jang, M. Beccani, J. Liang, S. Dixit, and R. Mangharam, "In-silico pre-clinical trials for implantable cardioverter defibrillators," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 169–172.

[12] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. CRC press Boca Raton, FL, 2014, vol. 2.

[13] R. D. Berger *et al.*, "The Rhythm ID Going Head to Head Trial (RIGHT): Design of a Randomized Trial Comparing Competitive Rhythm Discrimination Algorithms in Implantable Cardioverter Defibrillators," *Journal of Cardiovascular Electrophysiology*, vol. 17, no. 7, pp. 749–753, 2006.

[14] A. Fontenla, "Incidence of arrhythmias in a large cohort of patients with current implantable cardioverter-defibrillators in spain: results from the umbrella registry," *EP Europace*, vol. 18, no. 11, pp. 1726–1734, 2016.