ASME Letters in Dynamic Systems and Control
Online journal at:
https://asmedigitalcollection.asme.org/lettersdynsys

Ali Tivay
Mechanical Engineering,
University of Maryland,
2181 Glenn L. Martin Hall,
College Park, MD 20742;
Office of Science and Engineering Laboratories,
Center for Devices and Radiological Health,
U.S. Food and Drug Administration,
Silver Spring, MD 20903
e-mail: tivay@umd.edu

Ramin Bighamian
Office of Science and Engineering Laboratories,
Center for Devices and Radiological Health,
U.S. Food and Drug Administration,
10903 New Hampshire Ave,
Silver Spring, MD 20993
e-mail: ramin.bighamian@fda.hhs.gov

Jin-Oh Hahn[1]
Fellow ASME
Mechanical Engineering,
University of Maryland,
2181 Glenn L. Martin Hall,
College Park, MD 20742
e-mail: jhahn12@umd.edu

Christopher G. Scully
Office of Science and Engineering Laboratories,
Center for Devices and Radiological Health,
U.S. Food and Drug Administration,
10903 New Hampshire Ave,
Silver Spring, MD 20993
e-mail: christopher.scully@fda.hhs.gov

A Generative Approach to Testing the Performance of Physiological Control Algorithms

Physiological closed-loop control algorithms play an important role in the development of autonomous medical care systems, a promising area of research that has the potential to deliver healthcare therapies meeting each patient's specific needs. Computational approaches can support the evaluation of physiological closed-loop control algorithms considering various sources of patient variability that they may be presented with. In this article, we present a generative approach to testing the performance of physiological closed-loop control algorithms. This approach exploits a generative physiological model (which consists of stochastic and dynamic components that represent diverse physiological behaviors across a patient population) to generate a select group of virtual subjects. By testing a physiological closed-loop control algorithm against this select group, the approach estimates the distribution of relevant performance metrics in the represented population. We illustrate the promise of this approach by applying it to a practical case study on testing a closed-loop fluid resuscitation control algorithm designed for hemodynamic management. In this context, we show that the proposed approach can test the algorithm against virtual subjects equipped with a wide range of plausible physiological characteristics and behavior and that the test results can be used to estimate the distribution of relevant performance metrics in the represented population. In sum, the generative testing approach may offer a practical, efficient solution for conducting preclinical tests on physiological closed-loop control algorithms. [DOI: 10.1115/1.4065934]

Keywords: biological systems applications, biomedical systems, physiological closed-loop control, algorithm testing, hardware-in-the-loop testing

1 Introduction

Autonomous medical care is an emerging area of research that has the potential to provide the care that a patient needs while reducing clinician's workload. Central to this vision are the physiological closed-loop control algorithms, encompassing facets of perception, learning, and decision/control, that continuously monitor a patient's state and deploy treatments that are appropriate to the patient's condition. These algorithms are often constructed based on predefined assumptions pertaining to patient physiology and equipment characteristics. As a result, when the complexities of real-world situations diverge from these assumptions, the algorithms might face challenges in their operation. Hence, extensive testing under a wide range of scenarios can offer valuable insights into the behavior of these algorithms, e.g., by highlighting potential vulnerabilities and suggesting paths for further refinement.

Among available testing methods, in silico (using computer simulations) [1,2] and hardware-in-the-loop (HIL) [3,4] testing offer opportunities to gain valuable insights into the behavior of a given algorithm without imposing any direct risk to live animals or human subjects. Pivotal in these testing setups are patient physiology models capable of simulating various aspects of a patient's physiology that are relevant to the objectives of the test [5]. In recent years, there have been advances in mathematical models that represent patient physiology in a range of areas, including hemorrhage care [6], vasoplegia care [7], burn care [8,9], and diabetes care [10,11], to name a few. Concurrently, considerations and methodologies have emerged to enable credibility evaluation, verification, and validation for such mathematical models [5,12–15]. Furthermore, methods have been proposed to leverage these mathematical patient physiology models for in silico trials and/or to complement real-world trials with in silico results [1,2,12].

Generative physiological modeling is an emerging modeling concept that employs stochastic and dynamic components to emulate the physiological behaviors observed in a population of patients [6]. Generative physiological models can create cohorts of virtual patients, simulate the impact of physiological stimuli on these cohorts, and produce virtual datasets that mirror the spectrum

[1]Corresponding author.
Manuscript received February 27, 2024; final manuscript received June 29, 2024; published online August 2, 2024. Assoc. Editor: Reza Tafreshi.

of physiological responses anticipated in the population. In this article, we present a method to assess the performance of physiological closed-loop control algorithms using a generative physiological model. The presented method leverages the generative physiological model to create a curated group of virtual subjects. Then, the method offers estimates on the distribution of pertinent performance metrics in the population by examining the algorithm's performance in this curated group and analyzing the outcomes. To demonstrate this method, we apply it to a case study on HIL testing of a closed-loop fluid resuscitation control algorithm for hemodynamic management. In this context, we assess the method's ability to test a given control algorithm using virtual subjects that span varied physiological traits and whether the outcomes can estimate the distribution of relevant performance metrics in the represented population.

2 Generative Algorithm Testing

2.1 Generative Physiological Model. In this work, a generative physiological model [6] serves as the basis for testing the performance of physiological closed-loop control algorithms. The generative model used in this work consists of three main submodels:

$$\theta \sim \mathcal{G}(\phi_\mu, \phi_L) \tag{1}$$

$$x_k = \mathcal{H}(x_{k-1}, \hat{u}_{k-1}, \theta) \tag{2}$$

$$y_k = \mathcal{M}(x_k, \theta, n) \tag{3}$$

where $\mathcal{G}$ is the patient generator model and creates virtual patients by producing random Gaussian vectors θ representing patient characteristics and variations therein; $\mathcal{H}$ is the physiological dynamics model and generates changes in each virtual patient's state through time in response to inputs and therapies; and $\mathcal{M}$ is the physiological measurement model and represents the physiological measurement processes with their inherent imperfections. In this model, k is the time index, θ is the vector of patient characteristics, x_k is the vector of patient states, $\hat{u}_k$ is the vector of known inputs/therapies given to the patient, and y_k is the vector of generated measurements. Additionally, ϕ_μ , ϕ_L, and n denote the adjustable parameters of $\mathcal{G}$ and $\mathcal{M}$, where ϕ_μ and ϕ_L are the mean and the Cholesky decomposition of the covariance of the samples θ generated by the patient generator model $\mathcal{G}$, and n is the standard deviation (STD) of the noise considered by the measurement model. Given physiological data from a patient cohort, these adjustable parameters can be inferred via the method presented in our prior work [6]. This yields a generative physiological model capable of generating virtual datasets that share a similar distribution with the real data. Our objective in this work is to leverage this generative capability to test the performance of physiological closed-loop control algorithms.

2.2 Generative Algorithm Testing. In this section, we describe a conceptual framework for testing the performance of a given physiological closed-loop control algorithm using a generative physiological model. The control algorithm to be tested is formalized as

$$u_k = \mathcal{C}(*, \hat{y}_k) \tag{4}$$

where u_k is the control input (i.e., a command to administer therapy), $\hat{y}_k$ is the control output (i.e., relevant physiological variables measured from the patient), $\mathcal{C}$ is the controller, and $*$ indicates that the controller may be stateful. In real-world settings, the controller would interact with the patient through appropriate monitoring and therapy administration equipment. For testing purposes, we formalize such equipment as

$$\hat{y}_k = W_y(*, y_k) \tag{5}$$

$$\hat{u}_k = W_u(*, u_k) \tag{6}$$

where W_y denotes the monitoring equipment, W_u denotes the therapy administration equipment, and $*$ indicates that the equipment may be stateful. Eqs. (1)–(6) collectively outline a closed-loop testing setup built to assess the efficacy of a physiological closed-loop control algorithm by applying it to a virtual patient created by the generative physiological model.

Operating the testing setup through time yields the recording of time-series sequences, including inputs, outputs, and any relevant internal physiological variables simulated by the generative physiological model. These recordings serve as a basis to calculate performance metrics relevant to the objectives of the study, which can be formalized as follows:

$$m = \mathcal{P}(y_{1:T}, \hat{y}_{1:T}, u_{1:T}, \hat{u}_{1:T}, x_{1:T}) \tag{7}$$

where $\mathcal{P}$ denotes the performance evaluation process, m is a vector of performance metrics associated with the testing of the physiological closed-loop control algorithm on a virtual patient, and the subscript $1{:}T$ indicates the time-series nature of the variables where T is the end-of-test time index.

In pursuing the ultimate objective of generative algorithm testing, we generate a distribution of relevant performance metrics. This distribution represents the estimated values of the performance metrics in the population. Hence, it can highlight any potential deficiencies in the physiological closed-loop control algorithm that could arise due to variations in patient characteristics and/or nuances specific to the equipment. We formalize this distribution as follows:

$$p(m) = \mathbb{E}_{\theta\sim\mathcal{G}}[\delta(m|\theta)] \tag{8}$$

where δ is the Dirac delta distribution, $\mathbb{E}$ is the expectation operator, and $p(m)$ is the marginal probability density of the performance metrics. Direct numerical computation of this distribution requires the generation of a large number of virtual patients using the generative physiological model. The control algorithm is then tested to obtain the vector of performance metrics for each virtual patient. The individual results are then consolidated to form the distribution of the performance metrics. However, this approach could prove to be resource-intensive, particularly when implementing an HIL testing setup, or when the physiological model is computationally expensive to simulate. For example, it is not feasible to conduct HIL testing of a study protocol with a duration of several hours using a large number (tens of thousands [16]) of virtual patients. To mitigate this, in the following section, we present an efficient sampling approach aimed at estimating the distribution of the performance metrics with fewer tests.

2.3 Efficient Sampling for Performance Metric Estimation. To address the resource-intensive nature of the generative algorithm testing, we introduce an approach to generate a small select cohort of virtual patients and leverage this group to estimate the distribution of performance metrics. For this purpose, we focus on the parameters ϕ_μ and ϕ_L of the patient generator model in Eq. (1). Given ϕ_L, we compute its singular value decomposition (SVD) as

$$\phi_U \phi_S \phi_V^T = \text{SVD}(\phi_L) \tag{9}$$

Upon examining the results of the SVD, it is clear that $\phi_U\phi_S$ serves as an alternative square root of the covariance matrix $\phi_\Sigma = \phi_L\phi_L^T$. The columns of $\phi_U\phi_S$ point to the principal axes of the distribution of θ vectors produced by the patient generator model, arranged in the descending order of magnitude. Hence, these principal axes can be utilized to efficiently generate samples from the patient generator model by applying the unscented transformation formalism [17,18] as follows:

$$\theta_0 = \phi_\mu \tag{10}$$

$$\theta_{1:l} = \phi_\mu + \sqrt{l+\kappa}\,\phi_U\phi_S \tag{11}$$

$$\theta_{l+1:2l} = \phi_\mu - \sqrt{l+\kappa}\,\phi_U\phi_S \tag{12}$$

where θ_0, $\theta_{1:l}$, and $\theta_{l+1:2l}$ together denote $2l+1$ deterministic samples from the patient generator model, l denotes the size of θ, and κ is a tunable scaling factor. Given this array of virtual patients, we can execute the closed-loop testing setup in Eqs. (2)–(6) for each virtual patient. We can subsequently calculate a vector of performance metrics associated with each test as formalized in Eq. (7). The vectors generated by this process can be collectively utilized to estimate the mean $\bar{m}$ of the distribution pertaining to the performance metrics as follows:

$$\bar{m} \approx \frac{\kappa}{l+\kappa}[m|\theta_0] + \sum_{i=1}^{2l}\frac{1}{2(l+\kappa)}[m|\theta_i] \tag{13}$$

where $[m|\theta_i]$ denotes the value of m given the virtual patient characterized by θ_i. The covariance $\bar{M}$ of the distribution can likewise be estimated as follows:

$$\begin{aligned}\bar{M} \approx{}& \frac{\kappa}{l+\kappa}[m-\bar{m}|\theta_0][m-\bar{m}|\theta_0]^T \\ &+ \sum_{i=1}^{2l}\frac{1}{2(l+\kappa)}[m-\bar{m}|\theta_i][m-\bar{m}|\theta_i]^T\end{aligned} \tag{14}$$

The efficient sampling approach outlined above facilitates the estimation of performance metric means and covariances utilizing only $2l+1$ algorithm tests. However, depending on the particulars of the generative physiological model, this number may be reduced even further. In fact, prior empirical research suggests that a significant number of mathematical models in the fields of biology and physics possess parameterizations demonstrating sensitivities that span vast orders of magnitude [19]. This characteristic enables the formation of a "compressed" generative physiological model of patient physiology as described in [6], creating variations that align with the data distribution, while operating within a dimensionality less than l. This scenario implies that the elements of the diagonal matrix ϕ_S will span a broad range of values, where the first few diagonal elements are significantly larger than the subsequent elements. In such cases, it would be feasible to further approximate the metric distribution by generating deterministic samples associated with only the first l_c diagonal elements of ϕ_S, where $l_c < l$. This can be achieved by substituting only the first l_c columns of $\phi_U\phi_S$ in Eqs. (11) and (12) with $l \leftarrow l_c$ throughout the calculations in Eqs. (11)–(14). This adaptation allows us to approximately estimate the mean and covariance of the metric distribution using only $2l_c+1$ tests, thereby increasing the efficiency of the testing procedure.

3 Methods

In Sec. 2, we established a novel conceptual framework for efficient generative testing of physiological closed-loop control algorithms. In this section, we demonstrate an application of this framework to the testing of closed-loop fluid resuscitation control algorithms, which are control algorithms that adjust the administration of fluids to treat hypovolemia (i.e., low circulating blood volume) in critically ill patients [20]. This section provides further details on the methods used for this application.

3.1 Generative Physiological Model of Hemodynamic Responses to Hemorrhage and Fluid Resuscitation. For testing physiological closed-loop control algorithms in the proposed framework, a generative physiological model is a key element. In this application, we utilize a generative model from our previous work [6]. This generative model was created to mimic and capture the physiological responses to hemorrhage and fluid resuscitation in a population of large animal (sheep) subjects and possesses a structure conforming to Eqs. (1)–(3). To represent a wide spectrum of physiological characteristics across the subject population, we utilize the efficient sampling approach described in Sec. 2.3 and create nine virtual subjects. These virtual subjects were generated by letting $l_c = 4$ and $\kappa = 10$. Each one of these virtual subjects can accept time-series hemorrhage and fluid infusion rate signals as input to produce physiological responses to these signals, which encompass mean arterial blood pressure (MAP; the primary endpoint to be regulated by the fluid resuscitation control algorithm) and intermediate physiological variables such as arterial/venous blood volume, red blood cell volume, variables related to fluid exchange between the blood and interstitial fluid, and variables tied to the regulation of systemic vascular resistance and cardiac output. These virtual subjects therefore provide a platform to evaluate the performance of a given fluid resuscitation control algorithm in the presence of variations in subject characteristics.

3.2 Fluid Resuscitation Control Algorithm. Fluid resuscitation control algorithms are an emerging category of medical care algorithms intended to automate the treatment of hypovolemia in critically ill patients. A fluid resuscitation control algorithm operates by monitoring a patient's state via physiological measurements, e.g., MAP. Then, based on these measurements, it executes necessary therapies, e.g., crystalloid, or colloid infusions, so as to restore and maintain the patient's circulating blood volume. To demonstrate the generative algorithm testing approach proposed in this work, we employ a prototype closed-loop fluid resuscitation control algorithm. This algorithm consists of a two degrees-of-freedom proportional-integral (2DOF PI) control scheme to regulate the patient's MAP response, and an antiwindup mechanism to prevent excessive overshoot in the MAP response. Figure 1 provides a schematic representation of this control algorithm. In addition, the equations describing the algorithm are presented in discrete-time form below:

$$[z_c]_k = [z_c]_{k-1} + \delta t K_I(r_{k-1} - \hat{y}_{k-1}) + \delta t K_W[z_w]_{k-1} \tag{15}$$

$$[z_u]_k = K_P b(r_k - \hat{y}_0) - K_P(\hat{y}_k - \hat{y}_0) + [z_c]_k \tag{16}$$

$$[z_w]_k = \mathrm{sat}([z_u]_k) - [z_u]_k \tag{17}$$

$$u_k = \mathrm{sat}([z_u]_k) \tag{18}$$

In these equations, $\hat{y}_k$ denotes the control output, which is MAP; $\hat{y}_0$ denotes MAP at the time of controller engagement; r_k is the set-point, which is the target MAP; and u_k is the control input, which is the fluid infusion rate command. Moreover, K_p is the proportional gain, K_I is the integral gain, K_W is the antiwindup gain, b is the set-point weight on the proportional term, z_c is the integrator output, z_w is the antiwindup feedback, z_u is the unbounded infusion rate, and $\mathrm{sat}(\cdot)$ denotes the saturation function. As these equations are expressed in discrete-time form, δt is the sampling time, and $[\,\cdot\,]_k$ denotes the value of the variable inside brackets at time index k. To tune the parameters of the controller, we considered a second-order model of MAP response to fluid infusion as the plant. We then tuned the parameters of the 2-DOF PI controller to achieve a prespecified target value for the 0 dB gain crossover frequency pertaining to the open-loop frequency response. The closed-loop fluid resuscitation control algorithm, as described in Eqs. (15)–(18), serves as an example to illustrate our generative approach to testing the performance of physiological control algorithms.

3.3 Hardware-in-the-Loop Setup for Generative Testing. To account for the effects of equipment on the performance of the controller, we created an HIL setup for generative testing (Fig. 2), which is an extension of the HIL setup presented in our previous

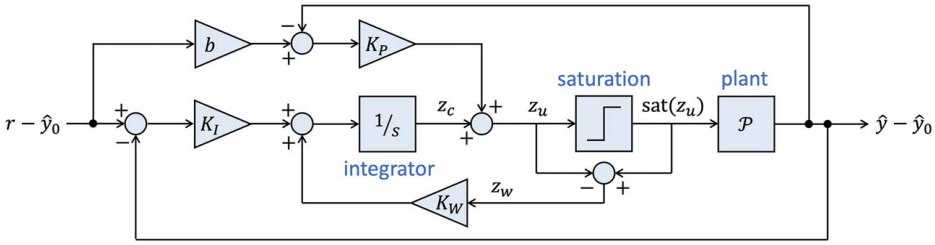

Fig. 1 Schematic representation of the fluid resuscitation control algorithm used to demonstrate the generative algorithm testing approach in this work

work [4]. This setup includes physical sensors and actuators in the control loop while simulating the mathematical models of patient physiology (i.e., virtual subjects) described in Sec. 3.1 as test subjects. Executing this setup involves a real-time iterative process, where: (i) the control algorithm suggests a fluid infusion rate; (ii) this rate is transmitted to an infusion pump; (iii) the actual amount of fluid infused by the pump is quantified and relayed to the virtual subject; (iv) the virtual subject is incrementally simulated using the actual quantity of the infused fluid; (v) the virtual subject's blood pressure waveform is generated and transferred to a physical pressure pulse generator and subsequently measured with a pressure transducer; and (vi) the transducer signal is processed by a patient monitor and channeled back to the control algorithm, thereby completing one iterative loop. By executing this iterative loop through time, we obtain a collection of time-series recordings from each virtual subject in the HIL setup. These recordings include relevant time-series data pertaining to both the control algorithm and the virtual subjects, which are then used to assess the performance of the control algorithm.

Within each HIL experiment, we adhered to a specific study protocol consisting of four stages: (i) first, each virtual subject experiences hemorrhage until its MAP decreases to 45 mmHg; (ii) a controlled hemorrhage is then applied, maintaining the subject's MAP at 45 mmHg for a duration of 15 min; (iii) after this maintenance period, the controller is engaged with a set-point of 70 mmHg. This allows us to test the controller's performance in restoring blood pressure to a target level; (iv) finally, 30 min after the controller's engagement, an additional episode of hemorrhage proportional to the subject's body weight is applied. This step further tests the controller's ability to maintain MAP against unknown disturbances.

4 Results and Discussion

4.1 Hardware-in-the-Loop Generative Testing. Figure 3 illustrates the data obtained from HIL controller tests executed on the virtual subjects, specifically focusing on the fluid infusion rate, hemorrhage rate, and MAP. Figure 4 presents the internal physiological variables corresponding to the virtual subjects presented in Fig. 3 (we refer the readers to Ref. [6] for a more detailed description of these physiological variables). These results show that the virtual subjects exhibit diverse physiological responses to the infusion rates applied by the controller. Specifically, the generated set includes virtual subjects that demand relatively modest volumes of fluid infusion to restore MAP to adequate levels, as well as subjects that require larger volumes, even saturating the infusion pump during certain periods. Further, the virtual subjects exhibit large variability even in terms of internal physiological variables, including arterial and venous blood volumes, red blood cell volume, parameters pertaining to the fluid exchange between the blood and interstitial fluid, and those related to the regulation of systemic vascular resistance and cardiac output. On the other hand, these responses, while varied, remain realistic and plausible.

Examination of Fig. 3 also suggests that the specific controller tested effectively manages the task of MAP regulation despite variations in the characteristics of the virtual subjects. It does so by applying a customized infusion rate profile to each virtual subject based on the subject's unique fluid infusion sensitivity, leading to a consistent MAP response across different virtual subjects. These results point to the potential of the proposed generative testing approach in efficiently evaluating a given control algorithm against a realistic spectrum of variations in subject characteristics.

4.2 Performance Metric Estimation. Building upon the approach outlined in Sec. 2.3, we utilize the collected results from nine individual HIL tests to estimate means and covariances associated with performance metrics in the population represented by the generative physiological model. To illustrate this, we use the procedure from Sec. 2.3 to estimate the means and covariances associated with five performance metrics relevant to fluid resuscitation control. To assess the controller's ability to regulate MAP to a target value when first engaged, we evaluate the rise time and the settling time of the MAP response (using data from 0–30 min in Fig. 3). To assess the controller's ability to reject the effect of disturbances (i.e., hemorrhage) on MAP while engaged, we evaluate the MAP drop during disturbance rejection and any subsequent overshoot associated with the rejection (using data from 30 min onwards in Fig. 3). To assess the controller's tendency to use up fluids toward its goals, we evaluate the total amount of fluid infused (using data from the entire timespan shown in Fig. 3).

Figure 5 visualizes the mean and STD estimates associated with the performance metrics described above. The MAP control rise time is estimated at 561 s with an STD of 34 s, and the MAP control settling time is estimated at 1316 s with an STD of 50 s. The MAP disturbance rejection overshoot is estimated at 2.3 mmHg with STD of 0.6 mmHg and the MAP drop during

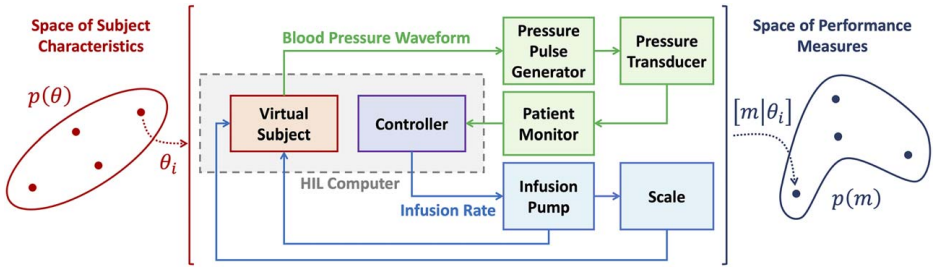

Fig. 2 Schematic representation of generative testing as applied to the hardware-in-the-loop testing of fluid resuscitation control algorithms

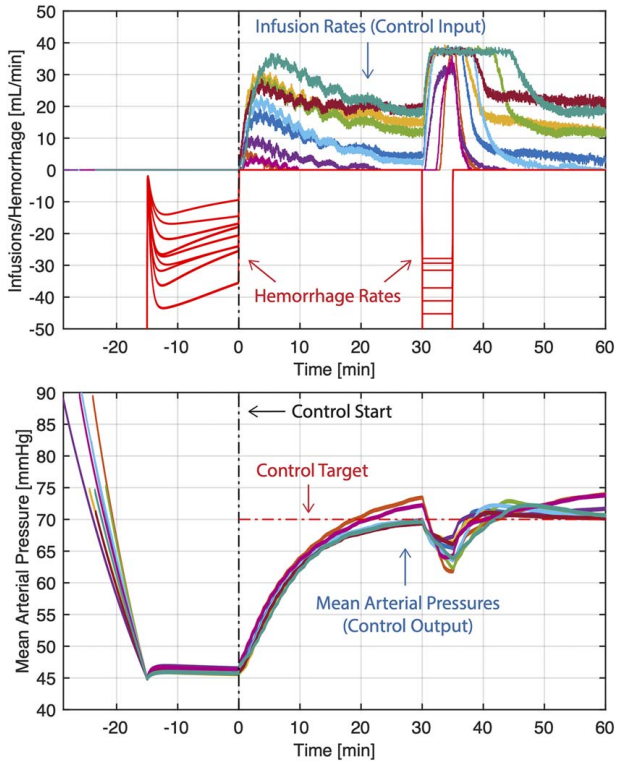

Fig. 3 HIL test results on virtual subjects, showcasing fluid infusion rate, hemorrhage rate, and MAP recordings

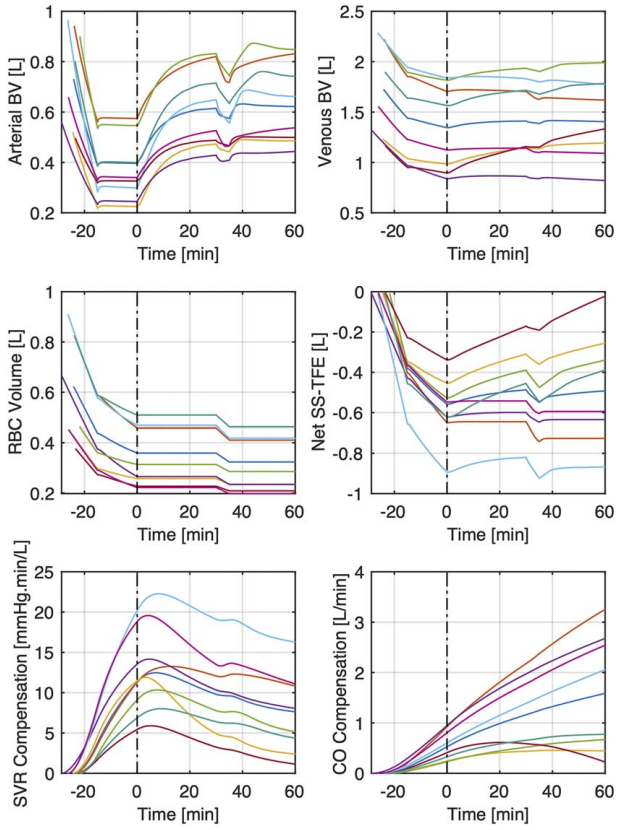

Fig. 4 Internal physiological variables of the virtual subjects depicted in Fig. 3 (BV: blood volume; RBC: red blood cells; SS-TFE: steady-state tissue fluid exchange; SVR: systemic vascular resistance; CO: cardiac output). SVR compensation and CO compensation denote autonomic baroreflex responses in SVR and CO in response to hemorrhage and fluid resuscitation.

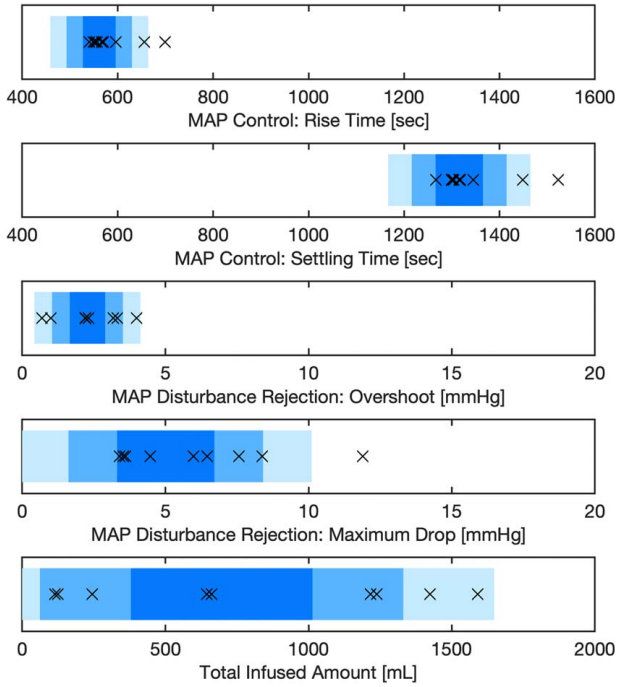

Fig. 5 Distribution estimates associated with the performance metrics of interest in fluid resuscitation control. Dark, medium, and light-shaded bands, respectively, visualize the $\mu \pm 1\sigma$, $\mu \pm 2\sigma$, and $\mu \pm 3\sigma$ bands of the estimated distribution, and points marked with "x" represent the metric values obtained from each HIL simulation.

disturbance rejection is estimated at 5 mmHg with STD of 1.7 mmHg. Finally, the total amount of fluid infused is estimated at 696 mL with an STD of 317 mL. As these estimates provide most-likely values and expected ranges for the performance metrics of interest, they underscore the benefits of the proposed generative testing framework in efficiently testing physiological closed-loop control algorithms and delivering estimates of metrics of interest in the population represented by the generative physiological model.

5 Conclusion

In this work, we introduced a method for testing physiological closed-loop control algorithms with a generative physiological model. Using a case study on HIL testing of closed-loop fluid resuscitation control algorithms, we showed the method's ability to test a given control algorithm against a diverse set of virtual subjects. The results from these tests provided insights into the algorithm's performance and offered an estimation of its performance across the population represented by the generative physiological model. As such, the proposed generative testing method offers a promising avenue for practical and efficient preclinical testing of physiological closed-loop control algorithms. Moving forward, it will be beneficial to prospectively validate the metric estimates produced by the method using new studies, demonstrate the effectiveness of this method in a wider range of algorithm testing scenarios, and expand the metric estimation techniques to a broader class of generative modeling frameworks.

Acknowledgment

The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services. This article reflects

the views of the author and should not be construed to represent FDA's views or policies.

Funding Data

- The U.S. National Science Foundation (Grant No. CMMI-1760817 with its INTERN Supplement).
- The U.S. National Science Foundation CAREER Award (Grant CNS-No. 1748762).

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

[1] Pappalardo, F., Russo, G., Tshinanu, F. M., and Viceconti, M., 2019, "In Silico Clinical Trials: Concepts and Early Adoptions," Briefings Bioinf., **20**(5), pp. 1699–1708.

[2] Viceconti, M., Emili, L., Afshari, P., Courcelles, E., Curreli, C., Famaey, N., Geris, L., et al., 2021, "Possible Contexts of Use for In Silico Trials Methodologies: A Consensus-Based Review," IEEE J. Biomed. Health Inform., **25**(10), pp. 3977–3982.

[3] Snider, E. J., Berard, D., Vega, S. J., Hernandez Torres, S. I., Avital, G., and Boice, E. N., 2022, "An Automated Hardware-in-Loop Testbed for Evaluating Hemorrhagic Shock Resuscitation Controllers," Bioengineering, **9**(8), p. 373.

[4] Mirinejad, H., Parvinian, B., Ricks, M., Zhang, Y., Weininger, S., Hahn, J. O., and Scully, C. G., 2020, "Evaluation of Fluid Resuscitation Control Algorithms Via a Hardware-in-the-Loop Test Bed," IEEE Trans. Biomed. Eng., **67**(2), pp. 471–481.

[5] Parvinian, B., Pathmanathan, P., Daluwatte, C., Yaghouby, F., Gray, R. A., Weininger, S., Morrison, T. M., and Scully, C. G., 2019, "Credibility Evidence for Computational Patient Models Used in the Development of Physiological Closed-Loop Controlled Devices for Critical Care Medicine," Front. Physiol., **10**(Mar.), p. 220.

[6] Tivay, A., Kramer, G. C., and Hahn, J. O., 2022, "Collective Variational Inference for Personalized and Generative Physiological Modeling: A Case Study on Hemorrhage Resuscitation," IEEE Trans. Biomed. Eng., **69**(2), pp. 666–677.

[7] Kao, Y. M., Sampson, C. M., Shah, S. A., Salsbury, J. R., Tivay, A., Bighamian, R., Scully, C. G., Kinsky, M., Kramer, G. C., and Hahn, J. O., 2023, "A Mathematical Model for Simulation of Vasoplegic Shock and Vasopressor Therapy," IEEE Trans. Biomed. Eng., **70**(5), pp. 1565–1574.

[8] Arabidarrehdor, G., Tivay, A., Meador, C., Kramer, G. C., Hahn, J. O., and Salinas, J., 2022, "Mathematical Modeling, In-Human Evaluation and Analysis of Volume Kinetics and Kidney Function After Burn Injury and Resuscitation," IEEE Trans. Biomed. Eng., **69**(1), pp. 366–376.

[9] Arabidarrehdor, G., Tivay, A., Bighamian, R., Meador, C., Kramer, G. C., Hahn, J. O., and Salinas, J., 2021, "Mathematical Model of Volume Kinetics and Renal Function After Burn Injury and Resuscitation," Burns, **47**(2), pp. 371–386.

[10] Dalla Man, C., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., and Cobelli, C., 2014, "The UVA/PADOVA Type 1 Diabetes Simulator: New Features," J Diabetes Sci Technol, **8**(1), pp. 26–34.

[11] Visentin, R., Campos-Náñez, E., Schiavon, M., Lv, D., Vettoretti, M., Breton, M., Kovatchev, B. P., Dalla Man, C., and Cobelli, C., 2018, "The UVA/Padova Type 1 Diabetes Simulator Goes From Single Meal to Single Day," J. Diabetes Sci. Technol., **12**(2), pp. 273–281.

[12] Viceconti, M., Juarez, M. A., Curreli, C., Pennisi, M., Russo, G., and Pappalardo, F., 2020, "Credibility of in Silico Trial Technologies—A Theoretical Framing," IEEE J. Biomed. Health Inform., **24**(1), pp. 4–13.

[13] Viceconti, M., Pappalardo, F., Rodriguez, B., Horner, M., Bischoff, J., and Musuamba Tshinanu, F., 2021, "In Silico Trials: Verification, Validation and Uncertainty Quantification of Predictive Models Used in the Regulatory Evaluation of Biomedical Products," Methods, **185**(1), pp. 120–127.

[14] Jang, K. J., Pant, Y. V., Zhang, B., Weimer, J., and Mangharam, R., 2019, "Robustness Evaluation of Computer-Aided Clinical Trials for Medical Devices," ICCPS 2019—Proceedings of the 2019 ACM/IEEE International Conference on Cyber-Physical Systems, Montreal, Quebec, Canada, Apr. 16–18, pp. 163–173.

[15] Bighamian, R., Hahn, J. O., Kramer, G., and Scully, C., 2021, "Accuracy Assessment Methods for Physiological Model Selection Toward Evaluation of Closed-Loop Controlled Medical Devices," PLoS One, **16**(4), p. e0251001.

[16] Kanal, V., Pathmanathan, P., Hahn, J. O., Kramer, G., Scully, C., and Bighamian, R., 2022, "Development and Validation of a Mathematical Model of Heart Rate Response to Fluid Perturbation," Sci. Rep., **12**(1), pp. 1–16.

[17] Julier, S., Uhlmann, J., and Durrant-Whyte, H. F., 2000, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," IEEE Trans. Autom. Control, **45**(3), pp. 477–482.

[18] Wan, E. A., and van der Merwe, R., 2001, "The Unscented Kalman Filter," *Kalman Filtering and Neural Networks*, S. Haykin, ed., John Wiley & Sons, Ltd, Hoboken, NJ, pp. 221–280.

[19] Transtrum, M. K., Machta, B. B., Brown, K. S., Daniels, B. C., Myers, C. R., and Sethna, J. P., 2015, "Perspective: Sloppiness and Emergent Theories in Physics, Biology, and Beyond," J. Chem. Phys., **143**(1), p. 010901.

[20] Avital, G., Snider, E. J., Berard, D., Vega, S. J., Hernandez Torres, S. I., Convertino, V. A., Salinas, J., and Boice, E. N., 2022, "Closed-Loop Controlled Fluid Administration Systems: A Comprehensive Scoping Review," J. Pers. Med., **12**(7), p. 1168.