Inference-Based Subject Atypicality and Signal Quality Indicators for Physiological Data

Ali Tivay
University of Maryland
College Park, MD, USA

George C. Kramer
University of Texas Medical Branch
Galveston, TX, USA

Jin-Oh Hahn
University of Maryland
College Park, MD, USA

ABSTRACT

Physiological measurements are an integral part of many established and emerging engineering and biomedical applications that involve physiological modeling, physiological state estimation, and physiological closed loop control. In practice, such measurements exhibit a large degree of variability, which is apparent at multiple levels, including disturbances acting on measured signals and unexpected physiological behavior in certain individuals. In this short paper, we present an inference-based approach to estimating the atypicality of an individual's physiological data both at the level of measurement and physiological behavior. For this purpose, we use data from a cohort of subjects to infer, simultaneously, model representations for measurement disturbances and atypicality of physiological behavior. Using a case study on hematocrit (HCT), cardiac output (CO), and mean arterial pressure (MAP) measurements in response to hemorrhage and colloid infusions, we discuss the merits of the presented approach in deriving reliable subject atypicality and signal quality indicators for physiological data.

CCS CONCEPTS

• **Applied computing** → **Engineering**; **Systems biology**.

KEYWORDS

Physiological Data, Signal Quality, Probabilistic Inference, Hematocrit, Cardiac Output, Arterial Pressure, Hemorrhage, Resuscitation.

ACM Reference Format:
Ali Tivay, George C. Kramer, and Jin-Oh Hahn. 2021. Inference-Based Subject Atypicality and Signal Quality Indicators for Physiological Data. In *Medical Cyber Physical Systems and Internet of Medical Things Workshop, May 18th, 2021.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/0000000.0000000

1 INTRODUCTION

Physiological measurements play a fundamental role in building our understanding of physiological dynamics in health and disease, and physiological feedback signals are an integral part of emerging technologies in physiological state estimation and physiological closed-loop control [1, 8, 13, 21, 25]. Naturally, real-world measurements are subject to noise, known artifacts, or unknown disturbances. This is even more prominent in physiological measurements, where (i) biological complexities and limitations of sensing often result in measurement signals with a high degree of variability in quality and information content, and (ii) a significant minority of individuals tend to show atypical physiological behavior in response to physiological stimuli. As a result, estimating the atypicality and quality of physiological signals, both at the level of measurement and physiological behavior, is a necessary step toward building reliable models and designing safe and effective state estimators and closed-loop controllers for physiological applications.

The design and estimation of signal quality indicators for physiological signals has received considerable attention from the research community. To this end, past researchers have typically divided their solution into two major steps involving the extraction of appropriate features, and leveraging those features to make decisions about signal quality [14, 15, 19]. Feature extraction is performed using a wide range of methods, including measuring aspects of the morphological shape of the signal [5, 9], matching the signal to a predefined set of templates [12, 26], or extracting the spectral and statistical characteristics of the signal [4, 23]. Decisions about signal quality are made using various discrimination techniques, including thresholds on known physiological features [2, 23], black-box machine learning techniques [16], and voting-based solutions that combine multiple discriminators [26]. Based on these approaches, many application-specific signal quality estimators have been proposed in the literature, especially for PPG [5, 9, 16], ECG [4, 19], EEG [23], and BCG/SCG [2, 26] signals. In addition, limited attempts have been made to propose more generalized algorithms and frameworks that can handle a wider category of signals [12, 20, 24].

Despite the sizable body of research on signal quality indicators for physiological signals, several important challenges still remain to be thoroughly resolved. First, most existing solutions are specialized to handle a specific type of signal (e.g. ECG), limiting their applicability to broader classes of physiological signals. Second, in addition to detecting and excluding poor-quality signals as a binary decision, it is desirable for the solutions to provide a full picture of the manner in which a signal is low-quality, which is often not provided in existing work. Third, in addition to the artifacts acting on physiological measurements, unexpected changes may arise in a signal due to atypical physiological characteristics and/or behavior in certain individuals, which is an important aspect that is rarely distinguished in existing solutions.

In a first step toward addressing these challenges, in this short paper, we present a potentially generalizable approach to estimating the atypicality of an individual's physiological data both at the level of measurement and physiological behavior. For this purpose, we

MCPS-2021, May 18, 2021, CPS-IoT Week

ACM ISBN 978-1-4503-XXXX-X/XX/XX.
https://doi.org/10.1145/0000000.0000000

cast the signal quality estimation problem as an inference problem on a generative model. The generative model is built to represent the population, the individuals, and the disturbances in a given physiological dataset. This approach enables us to use data from a cohort of subjects to simultaneously infer model representations for measurement disturbances and atypicality of physiological behavior. Using a case study on HCT, CO, and MAP measurements in response to hemorrhage and colloid infusions, we discuss the merits of the presented approach in deriving reliable subject atypicality and signal quality indicators for physiological data.

2 METHODS

In this section, we present our inference-based approach to subject atypicality and signal quality estimation for physiological data. The overarching idea is to (i) infer a generative model of the physiological data, and (ii) leverage the inference results to indicate the atypicality of the data, both in terms of measurement disturbances and the physiological behavior of the individuals. Details follow.

2.1 A Generative Model of Physiological Data

In this subsection, we describe our methodology for inferring a generative model for a given physiological dataset, with model elements that represent the population, the individuals, and the measurement disturbances present in the dataset.

For this purpose, we consider the following generative model of the physiological dataset:

$$\theta_i = \mu_\phi + \sigma_\phi \odot \epsilon \tag{1}$$

$$y_i = \mathcal{M}(\theta_i, u_i) + n_i \odot w \tag{2}$$

Equation (1) is a generative model of the population, where the output θ_i is a parameter vector representing the physiological characteristics of an individual (indexed by i), μ_ϕ is the mean of the generator, σ_ϕ is the standard deviation of the generator, and ϵ is a random vector drawn from a standard Gaussian distribution. Equation (2) is a generative model for physiological dynamics and measurements, where $\mathcal{M}$ denotes the physiological model (which may include mechanistic or black-box dynamics), u_i is a signal representing the known physiological stimuli received by the individual, and y_i denotes the model outputs corresponding to the individual. Furthermore, n_i is a parameter vector representing the output noise/disturbance characteristics, w denotes a white Gaussian noise signal, and $\odot$ denotes element-wise multiplication.

Given the generative model in (1)-(2), the objective is to identify the parameters of the model, maximizing the similarity between model generations and real physiological data. For this purpose, we start from computing the following posterior density [6] for individual characteristics:

$$P(\theta_i|y_i, u_i) = \frac{P_{n_i}(y_i|\theta_i, u_i)P_\phi(\theta_i)}{P(y_i|u_i)} \tag{3}$$

where $P_{n_i}(y_i|\theta_i, u_i)$ is the likelihood of the individual's physiological data with respect to the model in (2), which also depends on the disturbance characteristics n_i, and $P_\phi(\theta_i)$ is the density associated with the model in (1), which depends on $\phi = \{\mu_\phi, \sigma_\phi\}$. The denominator $P(y_i|u_i)$ denotes the model evidence. Next, we follow a variational approach [3, 11] to computing the posterior in (3). This approach approximates the true posterior $P(\theta_i|y_i, u_i)$ by searching over a family of approximate posteriors $Q_i(\theta_i)$. In this work, we use approximate posteriors of the diagonal Gaussian form:

$$Q_i(\theta_i) = \mathcal{N}(\mu_{\theta_i}, \mathrm{diag}(\sigma_{\theta_i})^2) \tag{4}$$

where μ_{θ_i} represents the most-likely value for the individual's physiological characteristics, and σ_{θ_i} represents the uncertainty associated with the individual's physiological characteristics. Having this family of approximate posteriors, it can be shown that the best approximate posterior is the one that maximizes the following evidence lower bound:

$$L_i = \log P(y_i|u_i) - D_{KL}\big[Q_i(\theta_i)||P(\theta_i|y_i, u_i)\big] \tag{5}$$

$$= E_{Q_i}\big[\log P_{n_i}(y_i|\theta_i, u_i) + \log P_\phi(\theta_i) - \log Q_i(\theta_i)\big] \tag{6}$$

where D_{KL} denotes the Kullback–Leibler divergence, E_{Q_i} denotes expectation with respect to samples from the approximate posterior, and L_i denotes the evidence lower bound (ELBO) associated with an individual. Having this individual-specific ELBO, the problem of inferring a generative model for the entire dataset can be summarized as the following optimization problem:

$$\mathbf{Q}^*,\ \mathbf{n}^*,\ \phi^* = \underset{\mathbf{Q},\,\mathbf{n},\,\phi}{\arg\max} \textstyle\sum_i L_i \tag{7}$$

where $\mathbf{Q}$ denotes the set of all Q_i's (each parameterized by μ_{θ_i} and σ_{θ_i}), $\mathbf{n}$ denotes the set of all n_i's, and the superscript $*$ denotes optimized parameters. The optimization problem in (7) is solved numerically by applying the "reparameterization trick" [11, 18] to the expectation operator in (6), and maximizing the objective using stochastic gradients of the terms inside the expectation [7, 10].

Overall, the presented optimization scheme uses data from a cohort of individuals to infer model representations for the severity of noises and artifacts acting on each signal (n_i^*'s), the likely physiological characteristics demonstrated by each individual (Q_i^*'s), and the occurrence density of each individual in the population (characterized by $P_{\phi^*}(\theta_i)$). Next, we leverage these results to define several useful signal quality indicators for the physiological data.

2.2 Inference-Based Subject Atypicality and Signal Quality Indicators

In this subsection, we derive several indicators, aimed at quantifying (i) atypical physiological behavior at the level of the individual, (ii) specific atypical physiological characteristics, and (iii) the noises and artifacts acting on each measurement signal.

To define an indicator that measures atypical physiological behavior in individual i, we consider our model representation of the population, which is characterized by $P_{\phi^*}(\theta_i)$, and measure the cumulative density of all the individuals that are more likely to occur in the population than individual i. For a Gaussian density, this quantity can be calculated as follows:

$$r_i = \sqrt{(\mu^*_{\theta_i} - \mu^*_\phi)^T\ \mathrm{diag}(\sigma^*_\phi)^{-2}(\mu^*_{\theta_i} - \mu^*_\phi)} \tag{8}$$

$$A_i = F(r_i, d_\theta) = \int_0^{r_i^2} \frac{t^{(d_\theta-2)/2}e^{-t/2}}{2^{d_\theta/2}\Gamma(d_\theta/2)}dt \tag{9}$$

where r_i is the (Mahalanobis) distance of the individual $\mu^*_{\theta_i}$ with respect to the population $P_{\phi^*}(\theta_i)$, d_θ is the dimension of θ_i, and A_i is the atypicality index for the individual. This atypicality index is

a number in the range $A_i \in [0, 1]$. In the extremes, if $A_i = 0$, no other individual is more likely to occur than individual i, indicating typical behavior. If $A_i = 1$, all other individuals are more likely to occur than individual i, indicating highly atypical behavior.

To define indicators for specific atypical physiological characteristics in individual i, we follow a procedure similar to the one above, but for each element j of the physiological characteristics vector $(\theta_i)_j$, which yields the following element-wise parameter atypicality index for the individual:

$$r_{ij} = [(\mu^*_{\theta_i})_j - (\mu^*_{\phi})_j]/(\sigma^*_{\phi})_j \tag{10}$$

$$A_{ij} = F(r_{ij}, 1) \tag{11}$$

The index A_{ij} is a number in the range $A_{ij} \in [0, 1]$, and measures the atypicality of the physiological characteristic j in individual i.

Finally, estimates for the severity of disturbances on individual i's measured signals are directly read from n^*_i, which contains standard deviation values for disturbances acting on each measured variable. Overall, the indices A_i, A_{ij}, and n^*_i can be used to assess the atypicality of an individual's physiological data both at the level of measurement and physiological behavior. In the next step, we present a case study to demonstrate the merits of this approach in subject atypicality and signal quality estimation.

2.3 Case Study on Hemorrhage Resuscitation

To demonstrate the performance of the subject atypicality and signal quality indicators, we apply the proposed approach to a case study on HCT, CO, and MAP measurements in individuals subjected to hemorrhage and colloid infusions. Toward this goal, in this section we present (i) an overview of the experimental protocol and the specifications of the physiological dataset, and (ii) an overview of the physiological model ($\mathcal{M}$ in (2)) used for this case study.

The studied physiological dataset is a dataset from our previous work [17], which contains $N = 5$ animal (sheep) subjects undergoing an initial large hemorrhage and two subsequent smaller hemorrhages. Each subject is then resuscitated with colloid infusions according to a rule-based algorithm. The total study duration for each subject is 180 minutes, and HCT, CO, and MAP measurements are performed at 5-minute intervals. Examples of the protocol and the measured values can be viewed in Figure 4.

The model $\mathcal{M}(\theta_i, u_i)$ for this case study is a mechanistic physiological model of the hemodynamic responses to hemorrhage and fluid resuscitation, described in our previous work [22]. In this model, the physiological characteristic vector θ_i consists of 14 physiological parameters (11 structural parameters and 3 initial conditions), and the input signal u_i consists of hemorrhage and infusion rates recorded in the experiments. Given these inputs, simulating the model produces predictions for HCT, CO, and MAP. Please refer to [22] for a detailed description of the model structure.

3 RESULTS AND DISCUSSION

This section presents the results of applying generative modeling, subject atypicality, and signal quality estimation (presented in Sections 2.1-2.2) to the hemorrhage resuscitation case study (presented in Section 2.3).

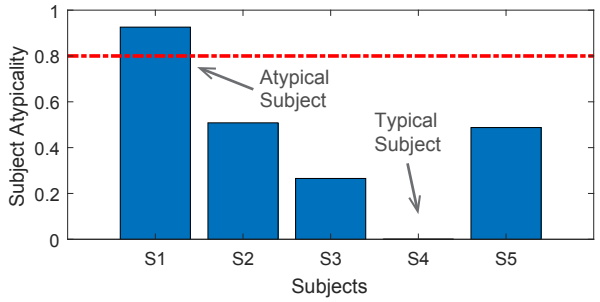

Figure 1: Individual-level atypicality index values (A_i) for the $N = 5$ subjects in the physiological dataset.

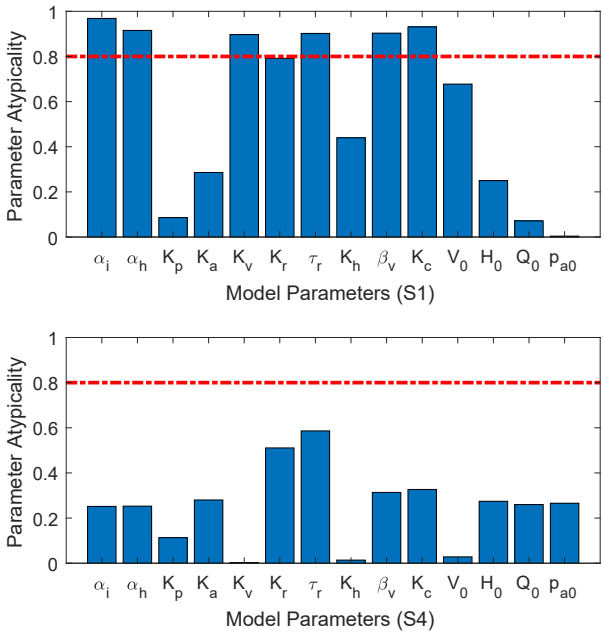

Figure 2: Parameter-level atypicality index values (A_{ij}) for the most atypical (S1, top panel) and the least atypical (S4, bottom panel) subjects in the dataset.

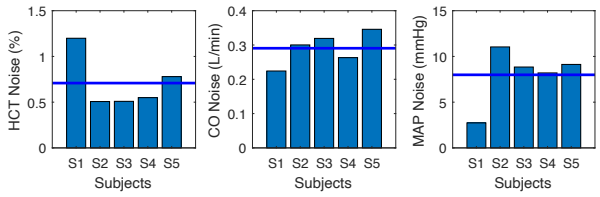

Figure 3: Signal-level disturbance severity estimates (n_i) for the $N = 5$ subjects in the physiological dataset.

Figure 1 shows the individual-level atypicality indices (A_i) for the subjects in the physiological dataset. The red line represents a threshold (at $A_i = 0.8$) that flags high subject atypicality. Subjects

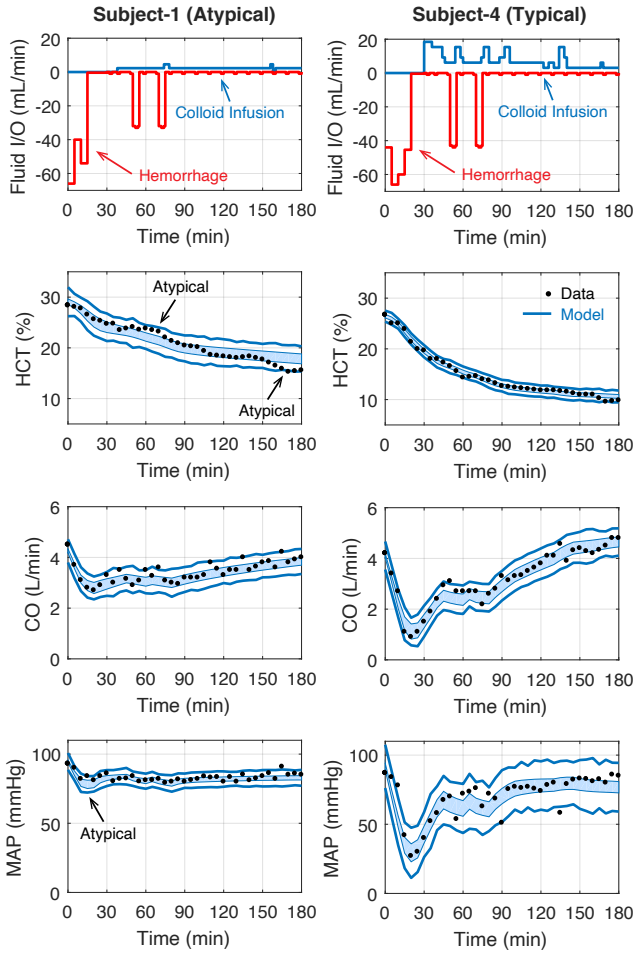

Figure 4: Physiological stimuli, measurements, and model responses for the most atypical (S1, left column) and the least atypical (S4, right column) subjects in the dataset.

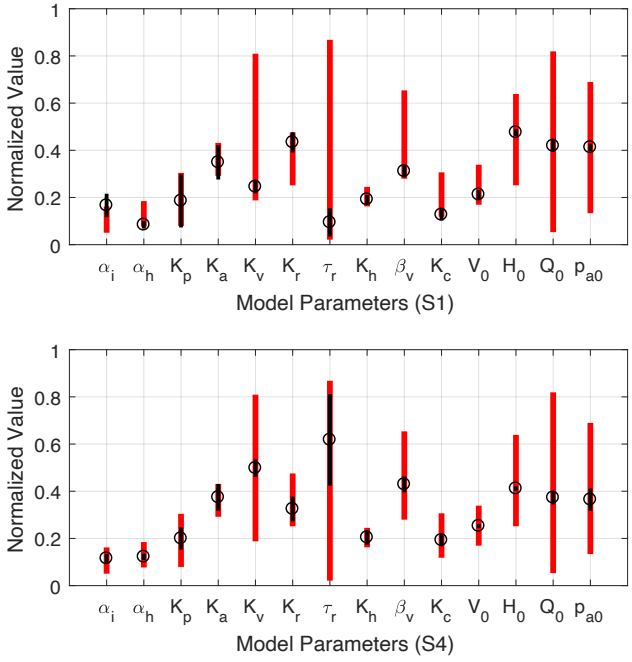

Figure 5: Inferred individual characteristics ($Q_i^*(\theta_i)$, black) vs. population characteristics ($P_{\phi^*}(\theta_i)$, red) for the most atypical (S1, top panel) and the least atypical (S4, bottom panel) subjects in the dataset. Vertical lines show 2σ confidence.

that pass this threshold are less likely to occur than at least 80% of the population. According to these indices, S1 is the most atypical subject, while S4 is the least atypical subject in the dataset. The difference in behavior between these two subjects can be observed visually by inspecting Figure 4. Typically, MAP and CO measurements should drop dramatically in response to a large hemorrhage, and HCT measurements should decrease in response to both hemorrhage and colloid infusions. However, according to the data shown in Figure 4, MAP in S1 does not drop dramatically in response to the large hemorrhage, and HCT in S1 shows two "bumps" that are not explainable by the hemorrhage and infusion profiles. These results suggest that the proposed method has correctly assigned a high atypicality index to S1.

Figure 2 shows the parameter-level atypicality indices (A_{ij}) for the most atypical (S1) and the least atypical (S4) subjects in the dataset. When inspecting these atypicality indices for S1, it can be observed that the subject shows highly abnormal characteristics in parameters associated with blood volume kinetics (e.g., α_i, α_h), and the regulation mechanisms that modulate cardiac output (e.g., β_v, K_c) and total peripheral resistance (e.g., τ_r, K_r). Further inspection of the parameter values in Figure 5 confirms that many of the S1's characteristics reside at the edges of the population characteristics. Overall, these results suggest that parameter-level atypicality indices (A_{ij}) may be used to gain further insight into the manner in which a subject shows atypical behavior.

Figure 3 shows the signal-level disturbance severity estimates (n_i) for the subjects in the physiological dataset. These estimates represent the standard deviation of the disturbances acting on each measured signal. For example, inspecting Figure 4 reveals that HCT measurements in S1 are affected by artifacts, while MAP measurements in S1 appear to have low noise. This is reflected in Figure 3 as a high index for HCT in S1 and a low index for MAP in S1. These results suggest that the signal-level disturbance severity indices (n_i) are useful representations of the noises/disturbances acting on the physiological measurements.

CONCLUSION

In this short paper, we proposed an inference-based approach to defining and quantifying the atypicality of an individual's physiological data both at the level of measurement and physiological behavior. In a case study on HCT, CO, and MAP measurements in response to hemorrhage and colloid infusion, we demonstrated that the proposed method can be utilized to obtain individual-level,

parameter-level, and signal-level atypicality indicators for physiological data. Given the promise of these initial results, future efforts should be devoted to generalizing and assessing this inference-based approach in a wider class of physiological signals.

ACKNOWLEDGMENTS

Research supported by National Science Foundation CAREER Award (Grant No. 1748762), and CDMRP (Grant No. W81XWH-19-1-0322).

REFERENCES

[1] Ghazal Arabidarrehdor, Ali Tivay, Ramin Bighamian, Chris Meador, George C Kramer, Jin-Oh Hahn, and Jose Salinas. 2021. Mathematical model of volume kinetics and renal function after burn injury and resuscitation. *Burns* 47, 2 (2021), 371–386.

[2] A Ozan Bicen and Omer T Inan. 2018. A signal quality index for ballistocardiogram recordings based on electrocardiogram RR intervals and matched filtering. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE, 145–148.

[3] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. 2017. Variational inference: A review for statisticians. *Journal of the American statistical Association* 112, 518 (2017), 859–877.

[4] GD Clifford, J Behar, Q Li, and Iead Rezek. 2012. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological measurement* 33, 9 (2012), 1419.

[5] Christoph Fischer, Benno Dömer, Thomas Wibmer, and Thomas Penzel. 2016. An algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *IEEE Journal of Biomedical and Health Informatics* 21, 2 (2016), 372–381.

[6] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian data analysis*. CRC press.

[7] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *The Journal of Machine Learning Research* 14, 1 (2013), 1303–1347.

[8] Xin Jin, Ramin Bighamian, and Jin-Oh Hahn. 2018. Development and in silico evaluation of a model-based closed-loop fluid resuscitation control algorithm. *IEEE Transactions on Biomedical Engineering* 66, 7 (2018), 1905–1914.

[9] Walter Karlen, K Kobayashi, J Mark Ansermino, and GA Dumont. 2012. Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation. *Physiological measurement* 33, 10 (2012), 1617.

[10] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[11] Diederik P Kingma, Max Welling, et al. 2019. An Introduction to Variational Autoencoders. *Foundations and Trends® in Machine Learning* 12, 4 (2019), 307–392.

[12] Qiao Li and Gari D Clifford. 2012. Dynamic time warping and machine learning for signal quality assessment of pulsatile signals. *Physiological measurement* 33, 9 (2012), 1491.

[13] Azin Mousavi, Ali Tivay, Barry Finegan, Michael Sean McMurtry, Ramakrishna Mukkamala, and Jin-Oh Hahn. 2019. Tapered vs. Uniform Tube-Load Modeling of Blood Pressure Wave Propagation in Human Aorta. *Frontiers in physiology* 10 (2019), 974.

[14] Shermeen Nizami, James R Green, and Carolyn McGregor. 2013. Implementation of artifact detection in critical care: A methodological review. *IEEE reviews in biomedical engineering* 6 (2013), 127–142.

[15] Christina Orphanidou. 2017. Signal quality assessment in physiological monitoring: state of the art and practical considerations. (2017).

[16] Tania Pereira, Cheng Ding, Kais Gadhoumi, Nate Tran, Rene A Colorado, Karl Meisel, and Xiao Hu. 2019. Deep learning approaches for plethysmography signal quality assessment in the presence of atrial fibrillation. *Physiological measurement* 40, 12 (2019), 125002.

[17] Abraham D Rafie, Paul A Rath, Michael W Michell, Robert A Kirschner, Donald J Deyo, Donald S Prough, James J Grady, and George C Kramer. 2004. Hypotensive resuscitation of multiple hemorrhages using crystalloid and colloids. *Shock* 22, 3 (2004), 262–269.

[18] Rajesh Ranganath, Sean Gerrish, and David Blei. 2014. Black box variational inference. In *Artificial Intelligence and Statistics*. PMLR, 814–822.

[19] Udit Satija, Barathram Ramkumar, and M Sabarimalai Manikandan. 2018. A review of signal processing techniques for electrocardiogram signal quality assessment. *IEEE reviews in biomedical engineering* 11 (2018), 36–52.

[20] Ikaro Silva, Joon Lee, and Roger G Mark. 2012. Signal quality estimation with multichannel adaptive filtering in intensive care settings. *IEEE Transactions on Biomedical Engineering* 59, 9 (2012), 2476–2485.

[21] Ali Tivay, Xin Jin, Alex Kai-Yuan Lo, Christopher G Scully, and Jin-Oh Hahn. 2020. Practical Use of Regularization in Individualizing a Mathematical Model of Cardiovascular Hemodynamics Using Scarce Data. *Frontiers in Physiology* 11 (2020).

[22] Ali Tivay, George C Kramer, and Jin-Oh Hahn. 2020. Virtual Patient Generation using Physiological Models through a Compressed Latent Parameterization. In *2020 American Control Conference (ACC)*. IEEE, 1335–1340.

[23] Jolanda Witteveen, Paruthi Pradhapan, and Vojkan Mihajlović. 2019. Comparison of a pragmatic and regression approach for wearable EEG signal quality assessment. *IEEE journal of biomedical and health informatics* 24, 3 (2019), 735–746.

[24] Farid Yaghouby, Chathuri Daluwatte, and Christopher G Scully. 2016. Variability analysis for noisy physiological signals: a simulation study. In *2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 1–5.

[25] Peyman Yousefian, Sungtae Shin, Azin Sadat Mousavi, Ali Tivay, Chang-Sei Kim, Ramakrishna Mukkamala, Dae-Geun Jang, Byung Hoon Ko, Jongwook Lee, Ui-Kun Kwon, et al. 2020. Pulse transit time-pulse wave analysis fusion based on wearable wrist ballistocardiogram for cuff-less blood pressure trend tracking. *IEEE Access* 8 (2020), 138077–138087.

[26] Jonathan Zia, Jacob Kimball, Sinan Hersek, Md Mobashir Hasan Shandhi, Beren Semiz, and Omer T Inan. 2019. A unified framework for quality indexing and classification of seismocardiogram signals. *IEEE journal of biomedical and health informatics* 24, 4 (2019), 1080–1092.