2020

Model-based machine learning to identify clinical relevance in a high-resolution simulation of sepsis and trauma

Zachary H. Silberman MD
UVM Larner College of Medicine, zachary.silberman@med.uvm.edu

Robert Chase Cockrell PHD
UVM Larner College of Medicine, robert.cockrell@med.uvm.edu

Gary An MD
UVM Larner College of Medicine, gary.an@med.uvm.edu

Title

"Model-based machine learning to identify clinical relevance in a high-resolution simulation of sepsis and trauma"

Introduction

Sepsis is a devastating and costly disease. According to the CDC, 1.7 million US adults develop sepsis, nearly 270,000 Americans die as a result of sepsis every year, and one in three patients who die in a hospital have sepsis [1,2]. It is a complicated process which represents the summation of varied host immune responses in a clinical and physiological diagnosis. Although improvements in care processes, such as the Surviving Sepsis Guidelines, have led to a significant reduction in mortality in the last 20 years[3], therapeutic options are limited to variations of antibiotics and physiologic support which date back nearly 25 years. Despite extensive research using current investigatory strategies, there is no current mediator-directed therapy, nor a biomarker panel able to categorize disease severity or reliably predict outcome. As such, one must ask, are these strategies capable of revealing the connection between basic science and potential therapeutics in a system as complicated as sepsis?

By looking at other fields of study, we find a possible answer: computational/simulation models as proxy systems. In systems engineering and other fields, these simplified systems advance understanding of their real world counterparts by providing discrete and knowable boundary conditions. Advances in computing power have enabled the development and testing of dynamic computational and mathematical models of acute systemic inflammation and sepsis. While these models are currently quite distant from direct clinical translation, they are able to provide critical insights into the dynamics of sepsis and can serve to frame the scope and direction of both basic and clinical research. Specifically, agent-based models (ABMs) are particularly well suited to translating basic science-derived mechanistic knowledge into heterogeneous clinical settings[4]. ABMs treat a particular system as populations of interacting components that are governed by rules for their behavior; the aggregate behavior of a system is thus simulated by creating and running the interactions between computational objects corresponding to those components. This can lead to emergent properties of the system not apparent from the rules governing agent interactions. Even with simple rules, very complicated patterns and relationships can arise. Additionally, because ABMs are defined by the rules governing each agent rather than the aggregate behavior, they are considerably easier to implement than equation-based models, such as those based on differential equations. Furthermore, they are adaptable to changing condition and readily allow for incorporation of stochastic factors[5]. As such ABMs are well suited to representing biological systems, where cells form a natural agent/component level for biomedical ABMs.

However, as complex objects themselves, the execution of ABMs is very computationally intensive. As mentioned previously, aggregate behavior is not immediately apparent and ABMs are frequently stochastic[6]. Consequently, ABMs require multiple simulations for sufficient validity and description of behavior. As such there is a need to develop new analytical methods to more efficiently and effectively extract knowledge from ABMs. One particularly computationally intensive task is the parameterization and calibration of ABMs. This process involves adjusting the values that affect the various mechanistic rules within the ABM and are analogous to the “weights” that govern the model’s responsiveness to simulated perturbations.

A specific ABM which has been well described and characterized in the 20 years since it was created, the Innate Immune Response Agent Based Model (IIRABM), remains a useful tool for research and characterization of sepsis. The IIRABM simulates the overall dynamics of the acute inflammatory response, producing recognizable aggregate outcomes (healing, pro-inflammatory death, hypo-immune death, and overwhelming initial infection) that are a simplified representation of the human innate immune system. However, its direct clinical applicability remains distant, in part due to the intensive computing time required to characterize the model and response to interventions. It is not unique in that the parameterization of the model is particularly intensive. In previous work[7], it was shown that the IIRABM can be characterized based on its four initializing parameters: microbial invasiveness, toxigenesis, environmental toxicity, and patient resilience. While they have no direct clinical correlates, they, along with the simulated outcomes, represent the parameter space of the model. Within this space, there are regions for which simulated patients are likely to have one of the specific outcomes listed above. These regions, termed Probabilistic Basins of Attraction (PBoAs), allow directed investigation for specific sets of parameters of clinical interested. For example, the regions where all simulated patients die of overwhelming infection or all heal rapidly, are not clinically relevant as they do not represent clinically plausible behavior, whereas regions with outcome uncertainty produce diverse model dynamics and more accurately simulate possible real world scenarios.

Within these regions, additional information can be gleaned from examining the trajectories of in silico patients. Although the outcome of aggregate trajectories is uncertain, for individual "patients", there comes a point after which its outcome is certain, either death or healing. Prior to this point, the trajectory exhibits stochastic behavior and the zone where this occurs is termed the stochastic zone. The boundaries of this zone are clinically relevant as they represent the point of maximum injury while still being recoverable and the point of maximum health while still progressing to death. Additionally, these boundaries change dependent on their proximity to the PBoAs. Thus, defining these stochastic zones allows for characterization of the model's clinical behavior.

However, the boundaries of the stochastic region change with each parameter set and whenever interventions, such as potential treatments, are introduced into the model. At present, there is no way to define the boundaries of the stochastic region and thus the PBoA without running the model throughout the entire parameter space. This represents millions of simulated runs and significant computing time. Any increase in efficiency in predicting the boundaries of the stochastic zone based on location in parameter space would represent a significant decrease in time required to assess the model.

This process of predicting boundaries based on a starting set of values is essentially a task of regression. In its most basic sense, regression involves fitting inputs to outputs based on a rule or set of rules. The simplest form is linear regression where these inputs and outputs are related by a linear function f(x) = ax+b, choosing the parameters a and b to minimize some measure of error between true outputs and predicted outputs. Linear regression, while extremely useful, is inadequate for all but the simplest relationships between inputs and outputs. The basic function can be augmented by including polynomial terms, but these still do not fill the needs of sufficiently large or difficult parameter regression as they are still limited by only one layer of inter-relatability. To identify deep patterns in simulation data, machine learning methods become necessary tools for regression. Artificial neural networks, a type of machine learning, introduce multiple layers of modeling and weights within each layer. In a basic single layer ANN, we can define a function to produce output o_1 from inputs i_1, i_2 as $\mathrm{f}(\mathrm{i}_1,\ \mathrm{i}_2) = \mathrm{a}\left(\mathrm{w}_{1_1}\mathrm{i}_1 + \mathrm{w}_{1_2}\mathrm{i}_2\right)$ where w_{1_1}, w_{1_2} are weights that adjust the impact of $\mathrm{i}_1, \mathrm{i}_2$ respectively and $\mathrm{a}(\mathrm{x})$ is

an activation function allowing for nonlinearity necessary for binary classification or logistic regression. This basic ANN, excluding the $\mathrm{a}(\mathrm{x})$ function is essentially a multilinear regression model as it combines two linear regression models to produce a single output. At this point, however, it still is not sufficient for characterizing more complicated problems. The next step, therefore, is to add more complexity. If we call our initial function l_1 we can imagine adding similar functions with the same inputs, but with different weights and activation functions, which themselves become the inputs of another function, which then can predict the desired output. For example, we can define:

$$l_1 = \mathrm{a}_1(\mathrm{w}_{1_1}\mathrm{i}_1 + \mathrm{w}_{1_2}\mathrm{i}_2), \mathrm{l}_2 = \mathrm{a}_2(\mathrm{w}_{2_1}\mathrm{i}_1 + \mathrm{w}_{2_2}\mathrm{i}_2), l_3 = \mathrm{a}_3(\mathrm{w}_{3_1}\mathrm{i}_1 + \mathrm{w}_{3_2}\mathrm{i}_2)$$

Related by a final equation $p = b_1(v_1 l_1 + v_2 l_2 + v_3 l_3)$. The weights of these equations can be sequentially adjusted to produce the lowest measure of error between the output of p and the actual output o_1. Furthermore, the model can be made arbitrarily complex by introducing more layers and functions within each layer. This process is what makes neural networks so well suited to identifying patterns in vast amounts of data. They have been employed successfully to predict outcomes in multiple biologic systems[8,9].

In our project we applied an Artificial Neural Network to a previously generated comprehensive data set from the IIRABM, and successively reduced the amount of training data while attempting to maintain the ability of the ANN to accurately predict clinically relevant tipping points given a particular parameter set.

Methods

The PBoA data set was obtained from IIRABM parameter sweep data which explored 40 injury sizes and 800 parameter combinations, each with 100 stochastic replicates. PBoA boundaries were calculated by determining the greatest oxygen deficit an *in silico* simulation could reach and still heal and the lowest oxygen deficit an *in silico* patient could reach and still die for each parameter combination. Up to 90% of the data was used for training, as discussed in the results, and 10% for testing/validation. In order to regress the PBoA boundaries, we utilized a fully connected deep network which takes a 4-dimensional parameter vector as input, feeding into two fully connected layers with 256 and 128 nodes respectively, separated by a dropout layer [10] with a 20% dropout ration, and finally to a single output node. The loss metric used to train this algorithm is the Mean-Squared-Error (MSE). Prediction variance and error bars were calculated through stochastic variations to the dropout layer, as demonstrated with regards to Active Learning for regression in [11]

Results and Discussion:

Figure 1 shows the accuracy of the ANN in predicting the boundaries of the stochastic zone versus percent training set withheld. A maximal accuracy of 96% (SD 0.0026) is, predictably, obtained when evaluating the largest portion of the dataset at 90%. However, the model can predict within an acceptable degree of accuracy with inclusion of 20% of the training set (accuracy 87.30%, SD 0.570). Note that this accuracy is calculated as an internal metric of the KERAS structure in which the ANN is built.

These tipping points define the boundaries of stochastic behavior of the IIRABM for a given parameter set. They represent the most harm a simulated patient can incur while then going on to fully heal or the

healthiest a simulated patient can be while the proceeding to die. If a simulated trajectory crosses these tipping points, it fate becomes definite; they are the boundaries of the life and death attractors for a specific parameter set. Prior to crossing the tipping points, the behavior is stochastic. This was shown in previous work by halting and re-seeding the trajectory[7]. As a trajectory approached the boundary, a greater percentage of re-seeded trajectories converged on the respective fate, but some, still would diverge to the opposite fate.

To reiterate, this behavior is only observed in certain regions of the IIRABM's parameter space for which the model's behavior is unknown. For any parameter set, the trajectory will be mapped to either the healthy attractor or the death attractor. For specific parameters in this region of parameter space, the behavior is unpredictable. These Probablistic Basins of Attraction characterize the IIRABM's behavior and examination of their boundaries may provide insight to future targets for interventions and therapeutics. However, the degree of unpredictability changes as we move through parameter space. Certain combinations of parameters will be more likely to result in simulated death than health, but still exhibit stochastic behavior; in other words, the boundaries of the stochastic zone are different for each parameter set. Thus, characterizing the stochastic boundaries defines the PBoAs which themselves define the model's behavior.

However, our present work identifies these tipping points with complete knowledge of the solution. As such, we are not predicting the boundaries for the stochastic zone, but confirming whether they can be predicted with limited information. As such, this project answers the question, is it possible to predict the boundaries with limited information of parameter space? Although it is useful as a concept, other methods to confirm the ANN's effectiveness at predicting the boundaries are necessary when this is put into practice with novel simulations of the IIRABM. We must define a metric that can determine whether the NN is fully trained on a data set which is not representative of the entire IIRABM parameter space.

The NN's effectiveness is determined by how well it "fits" the data. For any neural network, there is a fine line between overfitting and underfitting to the data available. In underfitting, the NN is not given enough data to be able to predict the outcome. This can happen when the NN has not had enough repetitions to connect the data properly or if the NN does not have enough complexity to deal with the connections in the data. In the case of insufficient complexity, we can say the NN has high bias, essentially the model is too rigid in structure to adapt to the data. A simple example of this would be attempting to fit a linear model to a polynomial function. In overfitting, the NN has seen too much data and is no longer generalizable to other data. Essentially, the NN becomes fitted to the noise in the dataset, termed having high variance. In either of these circumstances, the NN is not able to predict outcomes with any degree of confidence. As such It is necessary determine when a NN is properly fit and trained on the data and able to make accurate predictions.

There is no exact measure of fitment. However, it can be graphically represented and evaluated using validation curves examples of which are shown in figure 2. These plots model error against iterations in both the training set, and a withheld validation set. As the model is trained, the error or loss in both the training set and testing set should decrease over time and eventually flatten. One can evaluate fitment by examining how or if these error decrease. In the case of a model with high bias, which underfits the data, the validation and training curves will fail to converge as seen in fig 2D. If the NN is underfit due to insufficient training data as opposed to high bias, the amount of loss will be seen decreasing throughout

the training as opposed to flattening. Alternatively, if the NN has high variance and overfits, the training and validation curves will initially converge, but the validation loss will then start to increase as seen in fig 2B.

Putting these concepts together, we envision a neural net augmented workflow as follows: simulations run on the IIRABM with novel interventions introduced at some time point, following intervention, batches of 500 parameter sets are simulated to completion, stochastic trajectory analysis is perform3ed nd passed through a neural net, a validation curve is generated based on these parameter and examined for fitment. If the neural network displays underfitting, then more of the parameter space is explored and simulated until the neural network is properly fit. At this point, we can definitively the boundaries of the stochastic zone are defined, thus allowing us to characterize the PBoAs and so the IIRABM following a supposed intervention.

To rephrase our premise, we start with complete knowledge of data for the entire population, our parameter space, and determined that sampling 20% of our population allows for the population data as a whole to be characterized with 90% accuracy. In other words, we determined the necessary n for which valid assumptions of the population can be made. Just as with its biologic correlate, this represents a significant increase in efficiency in output of the IIRABM. Although the neural network is computationally intensive, it is 100 times cheaper to re-train the neural network than to run a single simulation of one parameter set on the IIRABM, which requires 10 hours of CPU time. Thus, our present work represents an opportunity for a monumental increase in possible output and discovery possible with the IIRABM.

Figures

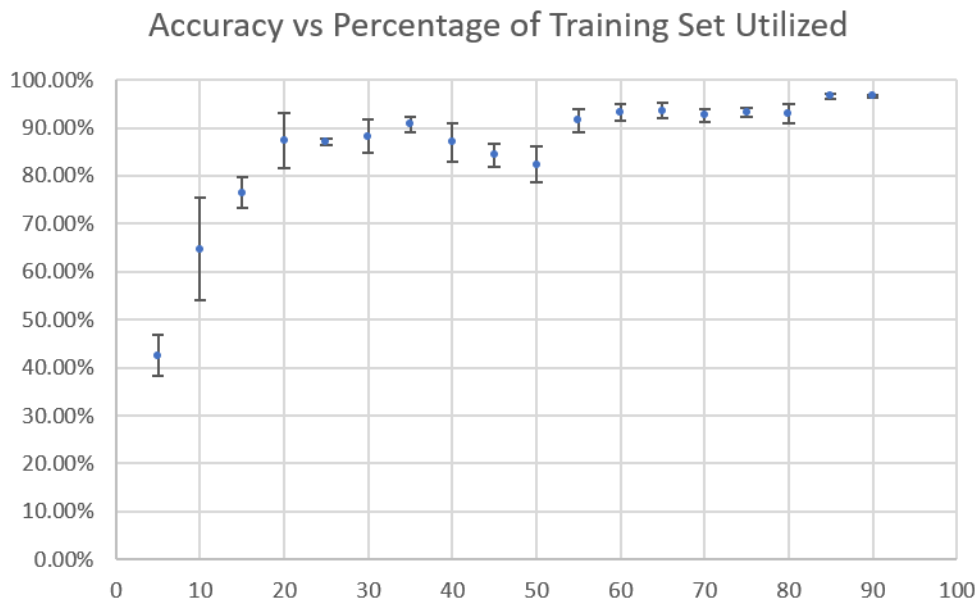

Figure 2 displays accuracy of the ANN in predicting boundaries of the stochastic zone based on percentage of total data used in the training set. Error bars are standard deviation.

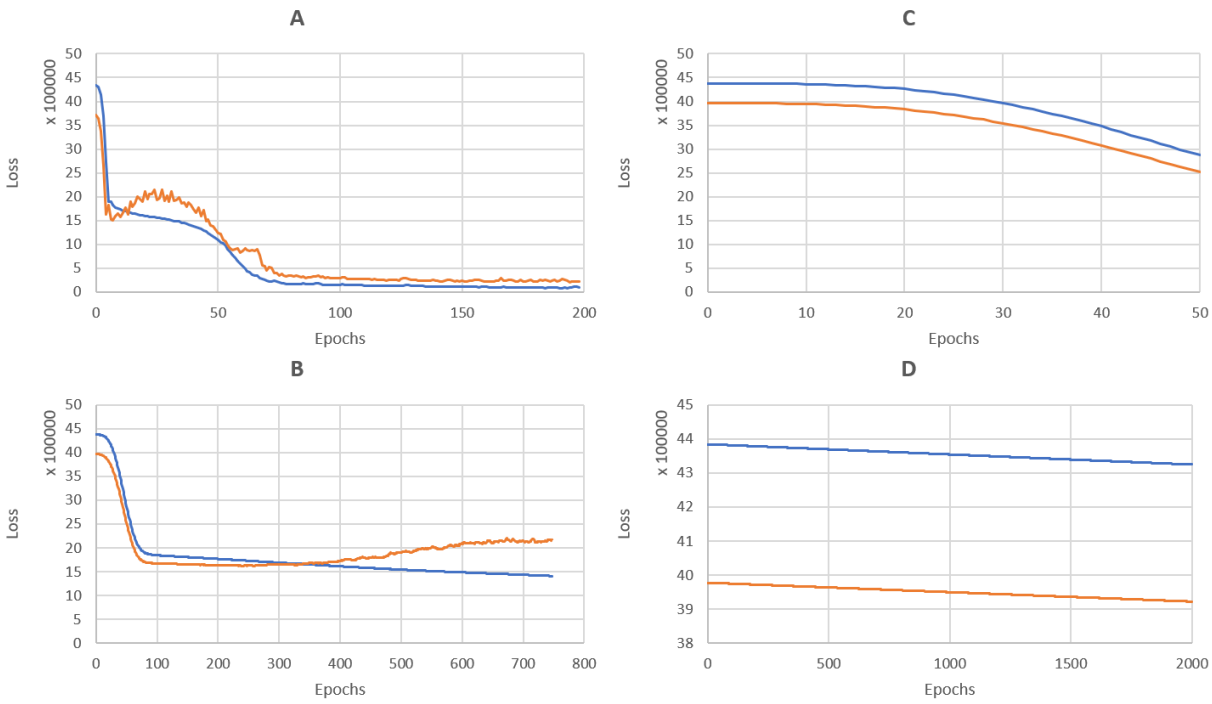

Figure 3 A-D. Examples of neural net validation curves plotting loss vs number of epochs. Training loss is shown in blue, validation set loss is shown in orange. All curves are a result of neural nets trained on the same data. 2A is an example of a well fit curve. the neural net here uses 4 layers of 256, 128, 128, and 64 nodes. note that both validation and training loss decrease significantly and converge as epochs progress indicating sufficient complexity of neural net. 2B demonstrates overfitting of a less complicated neural net with 2 layers of 16 and 4 nodes. Note that while the validation set and training set loss initially decrease, they begin to diverge after 400 epochs indicating overtraining, or high variance. 2C uses the same neural net as 2B, but training is ended after 50 epochs to show underfitting. Both validation set and training set loss are still decreasing before training has finished. 2D demonstrates underfitting due to high bias. the neural net consists of 1 layer with 1 node and is too rigid to predict the behavior of the model despite training to 2000 epochs.

References

1. Rhee C, Dantes R, Epstein L, et al. Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009-2014. *JAMA.* 2017;318(13):1241-1249.
2. Rhee C, Jones TM, Hamad Y, et al. Prevalence, Underlying Causes, and Preventability of Sepsis-Associated Mortality in US Acute Care Hospitals. *JAMA Network Open.* 2019;2(2):e187571-e187571.
3. Herrán-Monge R, Muriel-Bombín A, García-García MM, et al. Epidemiology and Changes in Mortality of Sepsis After the Implementation of Surviving Sepsis Campaign Guidelines. *Journal of Intensive Care Medicine.* 2017;34(9):740-750.
4. Ewald J, Sieber P, Garde R, Lang SN, Schuster S, Ibrahim B. Trends in mathematical modeling of host–pathogen interactions. *Cellular and Molecular Life Sciences.* 2020;77(3):467-480.
5. An G, Fitzpatrick BG, Christley S, et al. Optimization and Control of Agent-Based Models in Biology: A Perspective. *Bulletin of Mathematical Biology.* 2017;79(1):63-87.
6. Bauer AL, Beauchemin CAA, Perelson AS. Agent-based modeling of host–pathogen systems: The successes and challenges. *Information Sciences.* 2009;179(10):1379-1389.
7. Cockrell C, An G. Sepsis reconsidered: Identifying novel metrics for behavioral landscape characterization with a high-performance computing implementation of an agent-based model. *J Theor Biol.* 2017;430:157-168.
8. Subasi A, Erçelebi E. Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine.* 2005;78(2):87-99.
9. Heiat A. Comparison of artificial neural network and regression models for estimating software development effort. *Information and Software Technology.* 2002;44(15):911-922.
10. Baldi P, Sadowski PJ. Understanding dropout. Paper presented at: Advances in neural information processing systems2013.
11. Tsymbalov E, Panov M, Shapeev A. Dropout-based active learning for regression. Paper presented at: International Conference on Analysis of Images, Social Networks and Texts2018.