**2019 Multiscale Modeling Consortium Meeting - Translation and Dissemination (March 6-7, 2019)**

***Poster*** *Abstract Submission Form*

*\*Please submit to* [*https://msmmeeting.nibib.nih.gov/instructions-for-submitting-materials*](https://msmmeeting.nibib.nih.gov/instructions-for-submitting-materials) *by* ***February 1, 2018 (you may update your submission through this registration site after re-entering your login)***

*\*Save your abstract as “MSM PI Last Name \_ 2019 IMAG MSM Meeting Poster Abstract”*

**PI(s) of MSM U01: Gary An**

**Institution(s): University of Vermont**

**MSM U01 Grant Number: EB025825-01**

**Title of Grant:** Agent-based Models to address the Crisis of Reproducibility and

Precision Medicine

**Abstract Authors**

Gary An and Chase Cockrell

**Abstract Text**

***Introduction:*** *Sepsis in an inflammatory condition with a mortality rate of between 28%-50%. Current predictive models for sepsis generally use correlative methods, and as such are limited in their individual precision due to patient heterogeneity and data sparseness. Numerous mechanistic computational simulations of acute inflammation and sepsis have been utilized over the past two decades. These models have demonstrated that the sepsis population is much more heterogeneous than previously thought and this can be reflected by utilizing a range of multidimensional parameters that correlate to biologically plausible behaviors and phenotypes. Despite insights generated form these methods, there remain considerable challenges in the calibration, parameterization, and validation of the models. The description of the environment in which a biomedical simulation operates (model context) and parameterization of internal model rules (model content) requires the optimization (or near-optimization) of a large number of free-parameters; given the wide range of variable combinations, along with the intractability of* ***ab initio*** modeling techniques which could be used to constrain these combinations, an astronomical number of simulations would be required to achieve this goal. Active Learning (AL) can be combined with traditional machine learning techniques (in this case, artificial neural networks (ANN) and random forest (RF) models) to vastly reduce the size of this search space. ***Methods:*** *Billions of microbial sepsis patients were simulated representing a 28-day hospital course using a previously validated agent-based model (ABM) of sepsis, the Innate Immune Response Agent-Based Model (IIRABM), implemented on a Cray XE6 supercomputer. Contextual parameter space was examined regarding system response without antibiotics regarding the following parameters: cardio-respiratory-metabolic resilience; two properties of microbial virulence, invasiveness and toxigenesis; and degree of contamination from the environment. The model’s internal parameterization, which represents gene expression and associated cellular behaviors, was explored through the augmentation or inhibition of signaling pathways for 12 signaling mediators associated with inflammation and wound healing. We discretized the augmentation/inhibition space using 9 possible values for each cytokine pathway. We have employed the EMEWS (Extreme-scale Model Exploration with Swift) framework to implement a nested active learning approach in which the clinically relevant model environment space for a given internal model parameterization is mapped using either a small Artificial Neural Network (ANN) or Random Forest (RF) model. The outer AL level workflow is a larger ANN which uses active learning to efficiently regress the model coefficients for the lower level, as a function of the model’s internal parameterization.* ***Results:*** *A brute-force exploration of the IIRABM’s content and context would require approximately 3\*1012 simulations, and the end result would be a coarse representation of a continuous space.* *We have reduced the number of simulations required to efficiently map the clinically relevant parameter space of this model by approximately 95%. Additionally, we have shown that more complex models with a larger number of variables may expect further improvements in efficiency, thus effectively addressing the combinatorial explosion problem in computational model parameterization.* ***Conclusions:*** *Computational simulations make effective and appealing proxies for real world systems as they allow for experimentation and data collection at a scale that will never be feasible in vivo*; however, in order to provide clinical utility, *in silico* biomedical simulations must possess the same genetic variability as their *in vivo* correlates in the multi-phase clinical trial pipeline. Using EMEWS, we have demonstrated that the characterization of this genetic variability is now tractable using state-of-the-art machine learning methods in conjunction with high-performance computing.

*Please include in your abstract & poster how you are addressing the* ***10 Simple Rules for Credible Models***

|  |  |  |
| --- | --- | --- |
|  | | |
| 1 | Define context clearly | Develop and document the subject, application, purpose, and intended use(s) of the model or simulation |
| 2 | Use appropriate data | This work focuses more on modeling in general than on the validity of a single specific model |
| 3 | Evaluate within context | We are using machine learning techniques to find the boundaries of parameter/simulation space for which this model can be considered clinically relevant |
| 4 | List limitations explicitly | This work focuses more on modeling in general than on the validity of a single specific model, as such |
| 5 | Use version control | We are using the bitbucket repository for version control |
| 6 | Document adequately | We utilize commenting in our version control repository |
| 7 | Disseminate broadly | All codes are available for download from Bitbucket. Additionally, we are developing a lab webpage to store all of our model codes, papers, and supporting information in one place. |
| 8 | Get independent reviews | Our work is submitted to various respected journals, where it is peer-reviewed. |
| 9 | Test competing implementations | The LLNL and UVM factions of our collaboration maintain independent implementations of the IIRABM; for example, LLNL uses Boost to interact with the C++ model trough Python, while UVM uses Ctypes |
| 10 | Conform to standards | We perform a sufficient number of simulations to ensure that the results are repeatable, and any randomness/noise is well characterized |

***\*Save your abstract as “MSM PI Last Name \_ 2019 IMAG MSM Meeting Poster Abstract”***