Precision medicine as a control problem: Using simulation and deep reinforcement learning to discover adaptive, patient-specific multi-cytokine therapy for sepsis

D. M. Faissol, B. K. Petersen, C. P. Santiago, J. Yang (Georgia Tech), Chase Cockrell (Univ. of Vermont), and G. An (Univ. of Vermont)

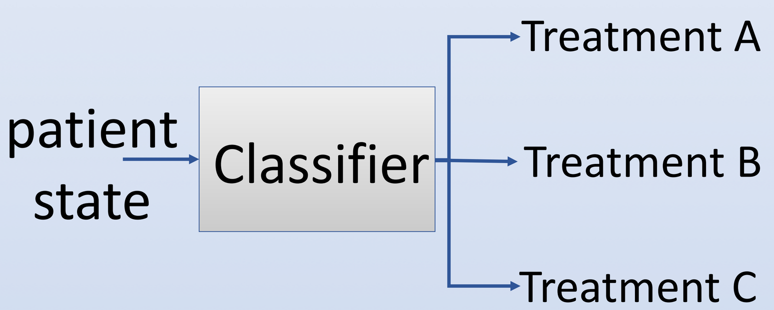
We are developing an approach for informing adaptive, patient-specific multi-drug therapeutic strategies with simulation and deep reinforcement learning. To demonstrate this approach, we are considering sepsis, a disease characterized by the dysregulation of immune response to infection or injury that results in millions of deaths annually and has no known cure. In this study, we attempt to discover an effective cytokine mediation treatment strategy for sepsis using a previously developed agent-based model that simulates the innate immune response to infection: the Innate Immune Response agent-based model (IIRABM).

Precision medicine as a control problem

Traditional precision medicine

Classify then *treat*

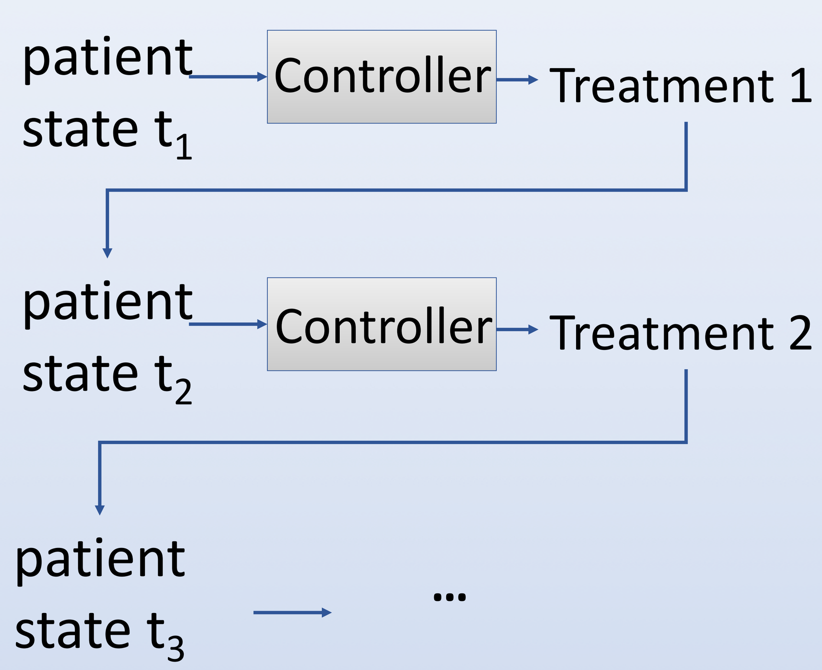
"…the ability to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment."
- National Research Council



- Viewed as a classification task
- Therapies are static and non-adaptive

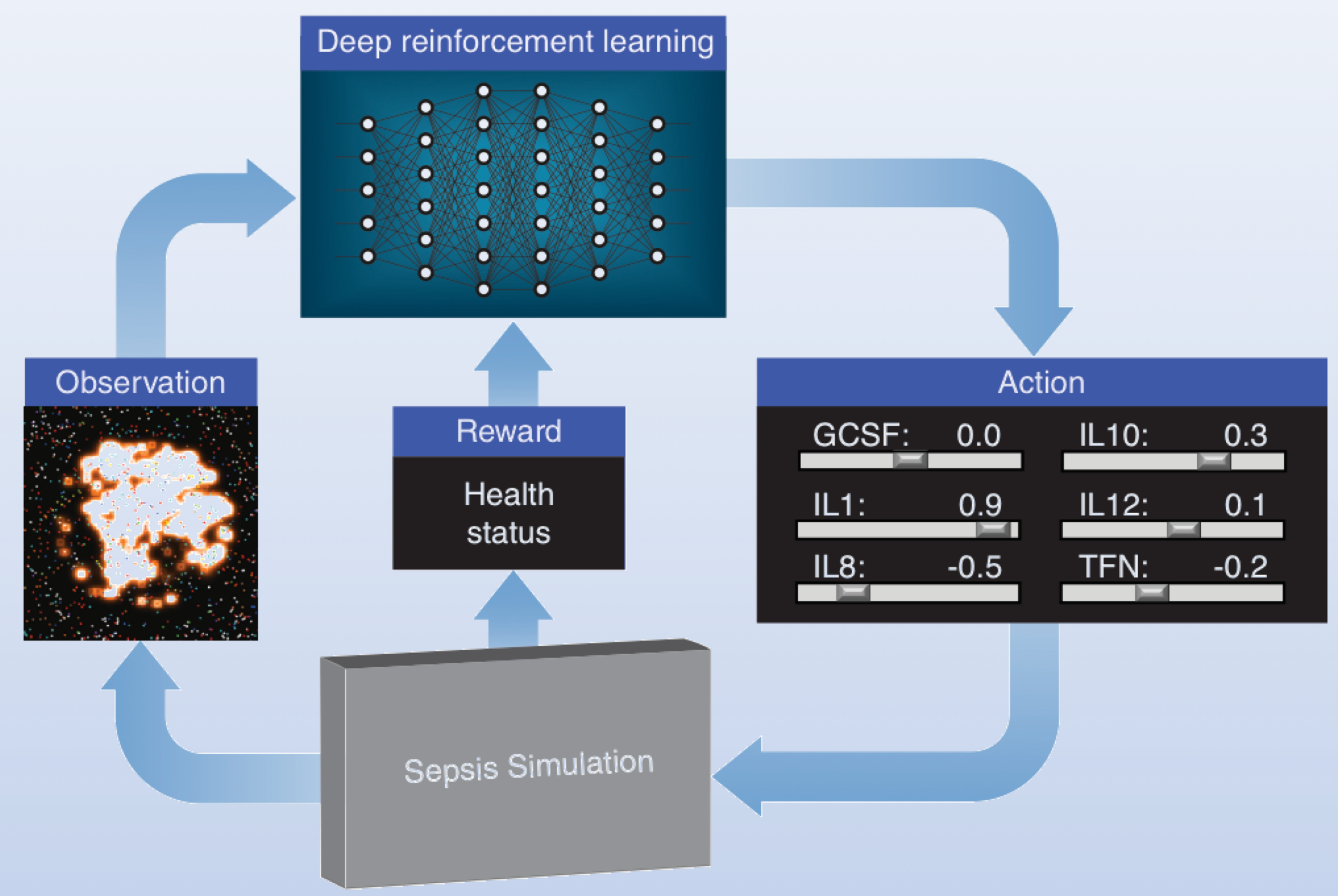
Proposed vision

Dynamic, feedback control



- Viewed as an optimal control task
- Therapies are dynamic and adaptive
  - Dependent upon patient trajectory

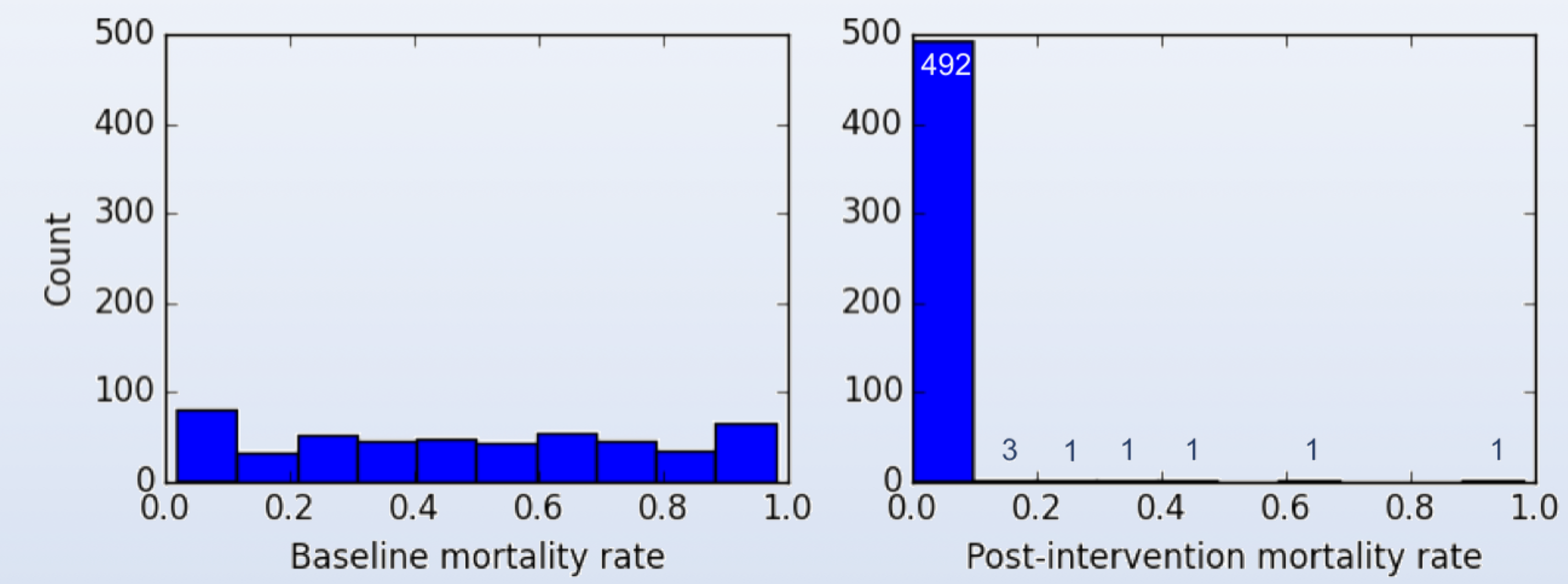
Learning adaptive therapies with deep reinforcement learning



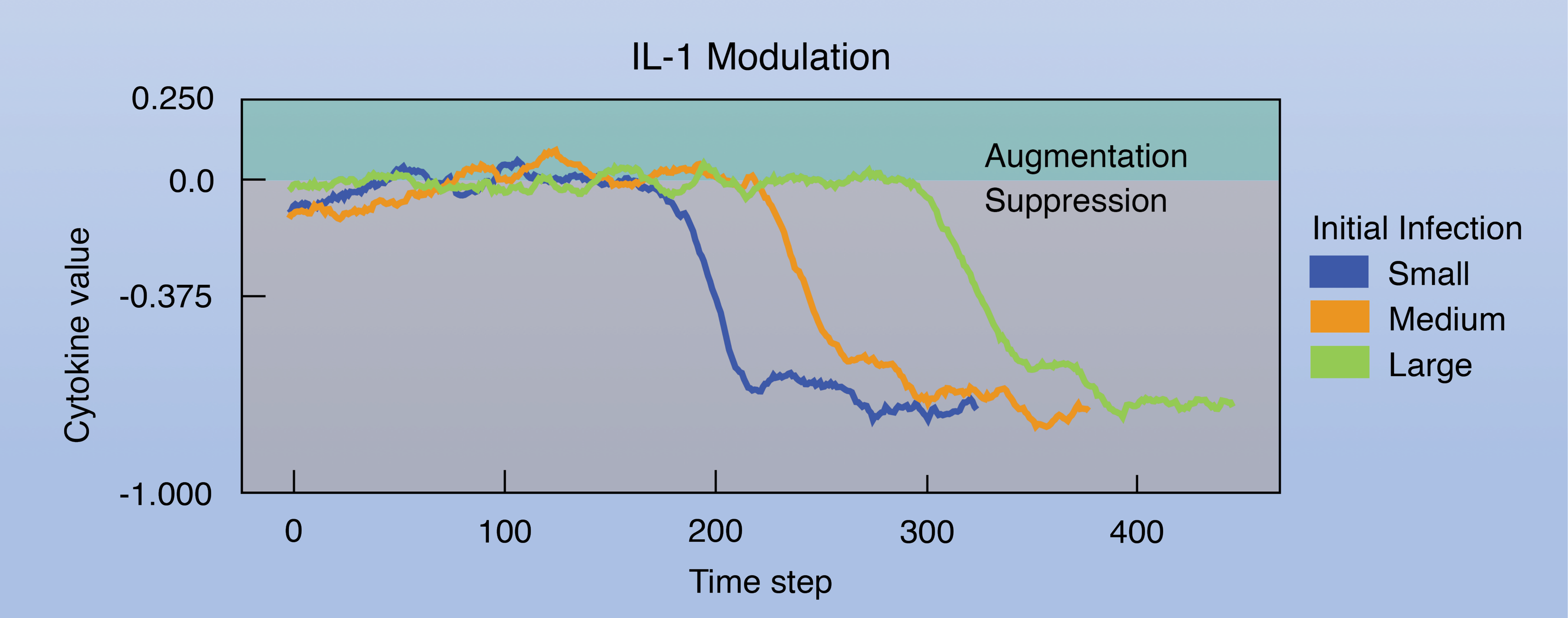
We use the IIRABM with a deep reinforcement learning algorithm to compute a treatment policy in which systemic patient measurements are used in a feedback loop to inform future treatment.

- Observation space: concentrations of the 12 cytokines, concentrations for 2 cytokine receptors, counts for 5 cell types, a measure of tissue damage, and a measure of infection
- Action space: increase or decrease any of the 12 cytokine concentrations
- Reward: determined by survival within the simulation, an penalty for taking actions, and a reward shaping term

Preliminary Results



- Histogram of mortality rates for 500 test patient parameterizations for antibiotics only (left) and deep reinforcement learning policy (right).
  - Average mortality rate with intervention is 0.8% vs 49% with antibiotics only
  - No simulated patients had higher mortality with intervention

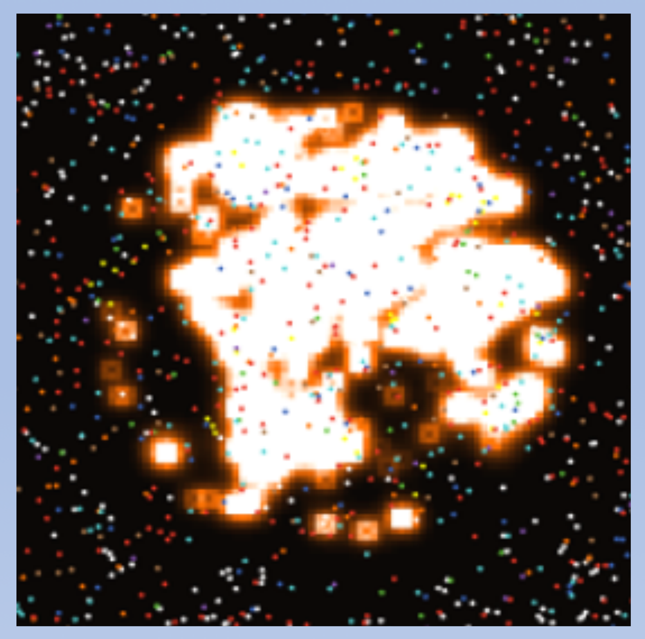


- Intuitive characteristics of the learned policy
  - IL-1 (pro-inflammatory) is unregulated early and suppressed late
  - Suppression comes later for patients with larger initial infections

Sepsis agent-based simulation model as a demonstration

Our sepsis agent-based model (ABM) captures:

- Five cell types (discrete agents)
- Twelve cytokine types (concentrations)
- Cell-activation states
- Infection level
- Tissue health (oxy) state
- Antibiotics



New, ongoing, research directions

More robust deep reinforcement learning approaches are needed

- Results generated by reinforcement learning are sensitive to the parameterization of the agent-based model
- Agent-based models of diseases have high uncertainties in parameter values
- Our new approach is
  1. Increase heterogeneity/uncertainty in model parameters of the IIRABM
  2. Constrain the parameter space by requiring the parameters to lead to plausible simulation output
  3. Search for cytokine mediation policies that are effective for all plausible parametrizations using more advanced DRL approaches

Lawrence Livermore National Laboratory

The University of Vermont

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.