What it will take to cross the Valley of Death:

Dealing with Biological Heterogeneity and Epistemic Uncertainty with Agent-based Modeling using an adaptation of the Principle of Maximal Entropy/Ignorance

The University of Vermont

LARNER COLLEGE OF MEDICINE

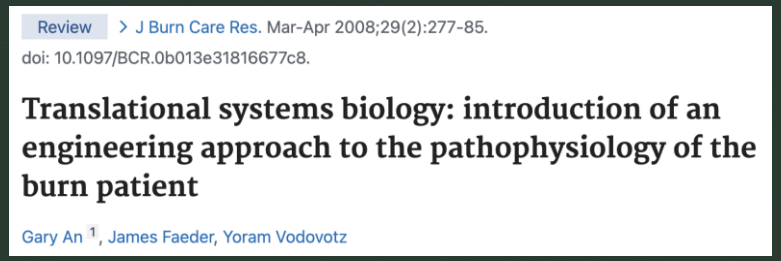
Gary An, MD, FACS
Department of Surgery
University of Vermont

U.S. Food and Drug Administration
FDA-MCERSI Workshop on QSP
May 11, 2023

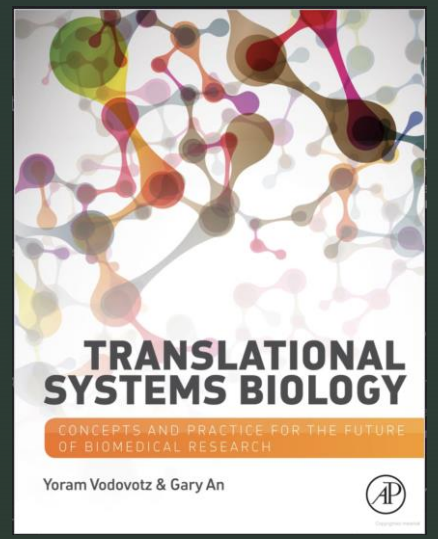
Translational Systems Biology

Name Motivated by the Zerhouni-Era NIH "Roadmap" 2003-5 => "Translational" Emphasis

2008



2015



Primary Features:

- Dynamic Computational Modeling to Accelerate Hypothesis Testing
- Simulate Clinical Contexts => In Silico Clinical Trials
- Use abstraction to identify conserved functions across biology (species, individuals, etc…)

Axioms of “True” Precision Medicine

- Axiom 1: Patient A is not the same as Patient B (Personalization)
- Axiom 2: Patient A at Time X is not the same as Patient A at Time Y (Precision)
- Axiom 3: The goal of medicine is to treat, prognosis is not enough (Treatment).
- ***Axiom 4: Precision medicine should find effective therapies for every patient and not only to identify groups of patients that respond to medicine (Inclusiveness)***

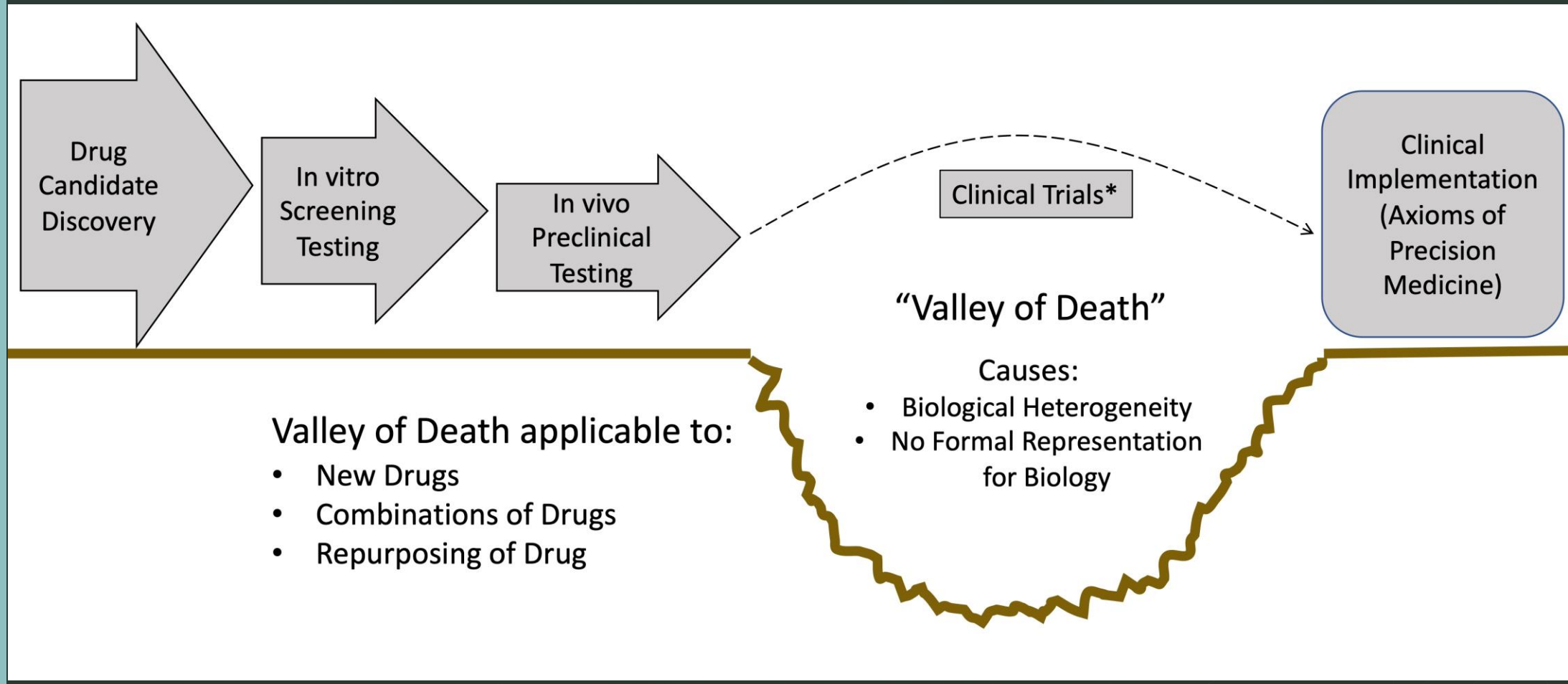
An and Day, “Precision Systems Medicine: A Control Discovery Problem”
Systems Medicine: Integrative, Qualitative and Computational Approaches, in press

Rare Diseases and Personalized/Precision Medicine

- Curse of Dimensionality:
  - # potential features + # points of interventions vs # patients with disease x their variance in manifestation => perpetual sparsity of data
- Is addressing “rare disease” same challenge as achieving personalized medicine?
- In terms of the degree of biological heterogeneity, everything is a ”rare” disease…

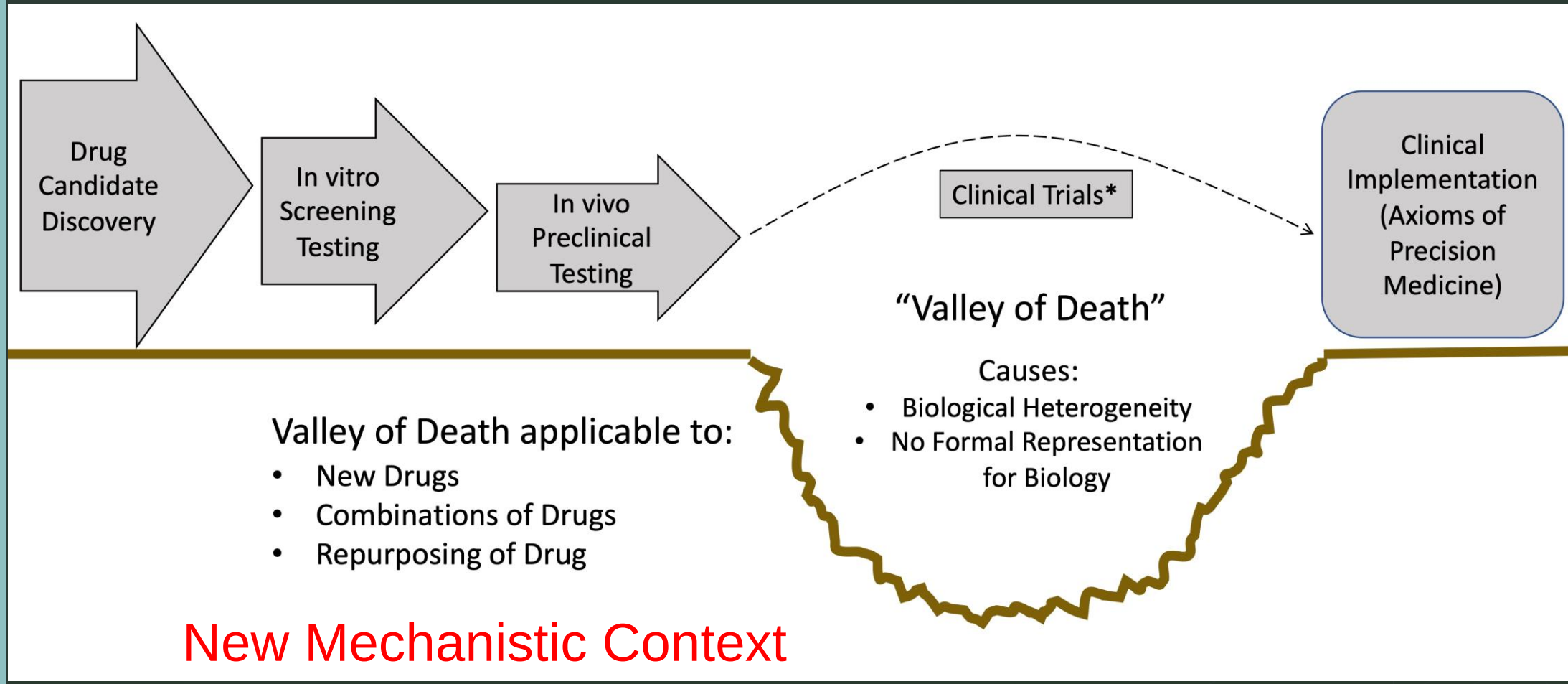
The Valley of Death

Most Impactful Failure Point in Drug Development



The Valley of Death

Most Impactful Failure Point in Drug Development



Reasons for the Valley of Death

- No “biological” natural laws/First Principles
- Biological heterogeneity => “Denominator Problem” of Perpetual Data Sparseness
- Epistemic Uncertainty:
  - Impossible to know everything
  - How to use partial information?

What other Sciences do = Engineering

- Engineering:
  - Formal Process for identifying solutions given a set of system constraints
  - Uses Math/Formal Representations
- Hard for what makes biology biology (i.e. behavior of cells)
  - No first principles/natural laws
  - Lack of “trustworthy” formal representations
  - Reverse engineering difficult due to non-path uniqueness and intrinsic heterogeneity (epistemic limits of current methods)

What other Sciences do = Engineering

- Engineering:
  - Formal Process for identifying solutions given a set of system constraints
  - Uses Math/Formal Representations
- Hard for what makes biology biology (i.e. behavior of cells)
  - No first principles/natural laws
  - Lack of “trustworthy” formal representations
  - Reverse engineering difficult due to non-path uniqueness and intrinsic heterogeneity (epistemic limits of current methods)
- How to use mathematical/computational/theoretical methods to overcome?

Medical Digital Twins and *In Silico* Trials: How to reconcile epistemic uncertainty and biological heterogeneity?

- *In Silico* trials should overcome barriers in Valley of Death. But…
- How to encompass human heterogeneity in a clinical population with incomplete knowledge (“good enough”)?
- How to capture what is similar but able to generate heterogeneity?
- How to deal with perpetual epistemic incompleteness?

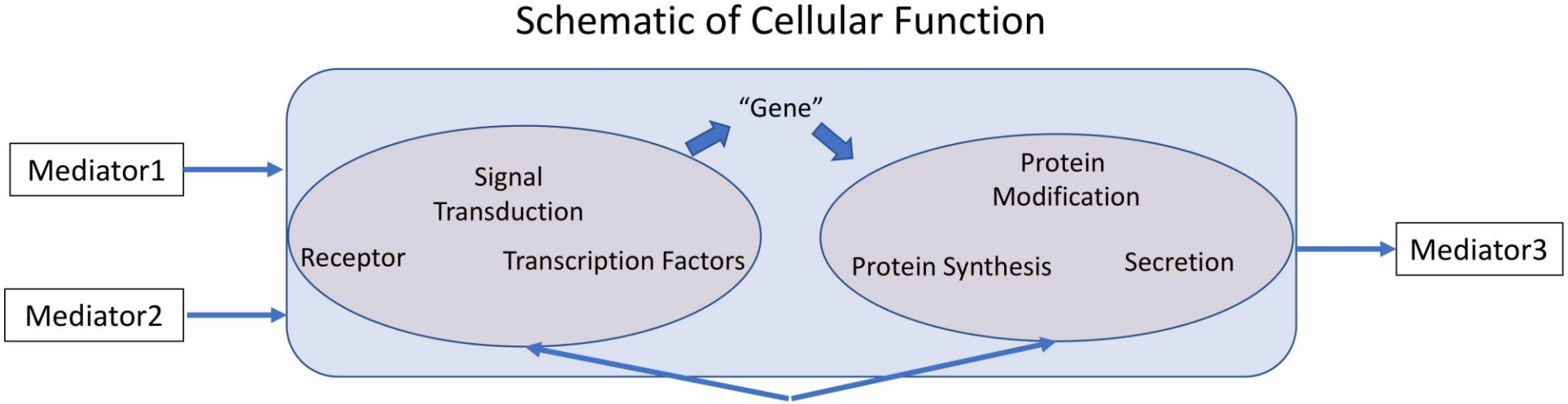
Dealing with Epistemic Uncertainty: Maximal Entropy Principle (MEP)

- Basis from Information Theory/Statistical Physics
  - Find least biased statistical model that reproduces a specific data set
  - Infers unknown causal interactions
  - Bio applications: gene/protein interactions, neural signal processing, ecosystem dynamics
- If you don't know the answer, it's better to have lots of possible answers than to be wrong*
- We extend MEP to the parameterization of mechanism-based models => **Embraces Heterogeneity**
- ***Goal of discovering robust control => guiding principle***

Context: Control Discovery

- Rational approach => requires a hypothesis of how the system works (mechanisms)
- Use minimal model “necessary” to establish control over desired features of the system
- Robust enough across biological/clinical heterogeneity and epistemic uncertainty
- Iterative refinement of underlying mechanistic model => more and “different” data

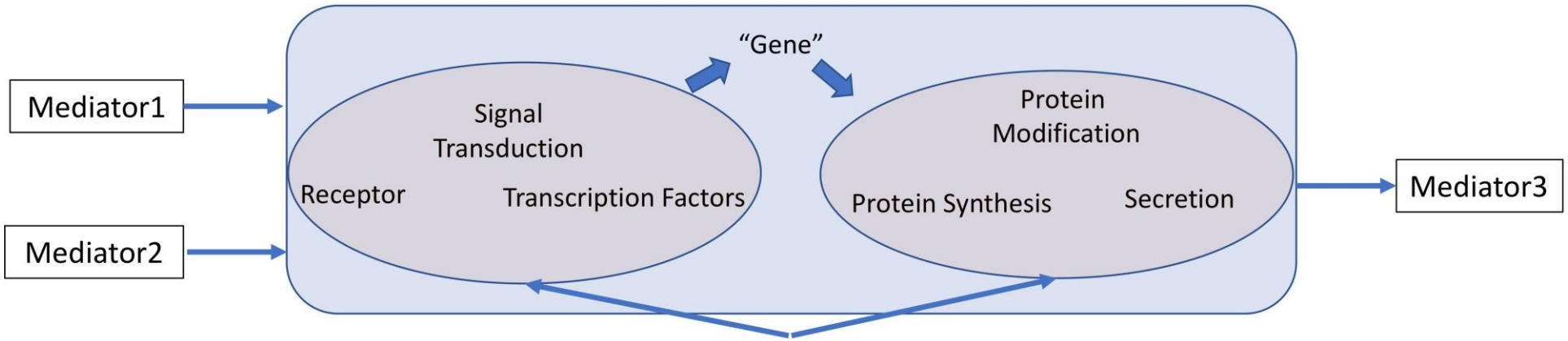
MEP applied to “Parameter” Landscape => Model Rule Matrix (MRM)



Form of Cellular Rule: aMediator1 + bMediator2 = Mediator3

Thus: ***a*** and ***b*** are rule parameters that represent “hidden” factors/controls/genes that affect the contribution of Mediator1 and Mediator2 to the production of Mediator3

MEP applied to “Parameter” Landscape => Model Rule Matrix (MRM)



All these factors “condensed” into par
into a representation of cell

Form of Cellular Rule: *a*Mec

Thus: ***a*** and ***b*** are rule parameters
that affect the contribution of Mec

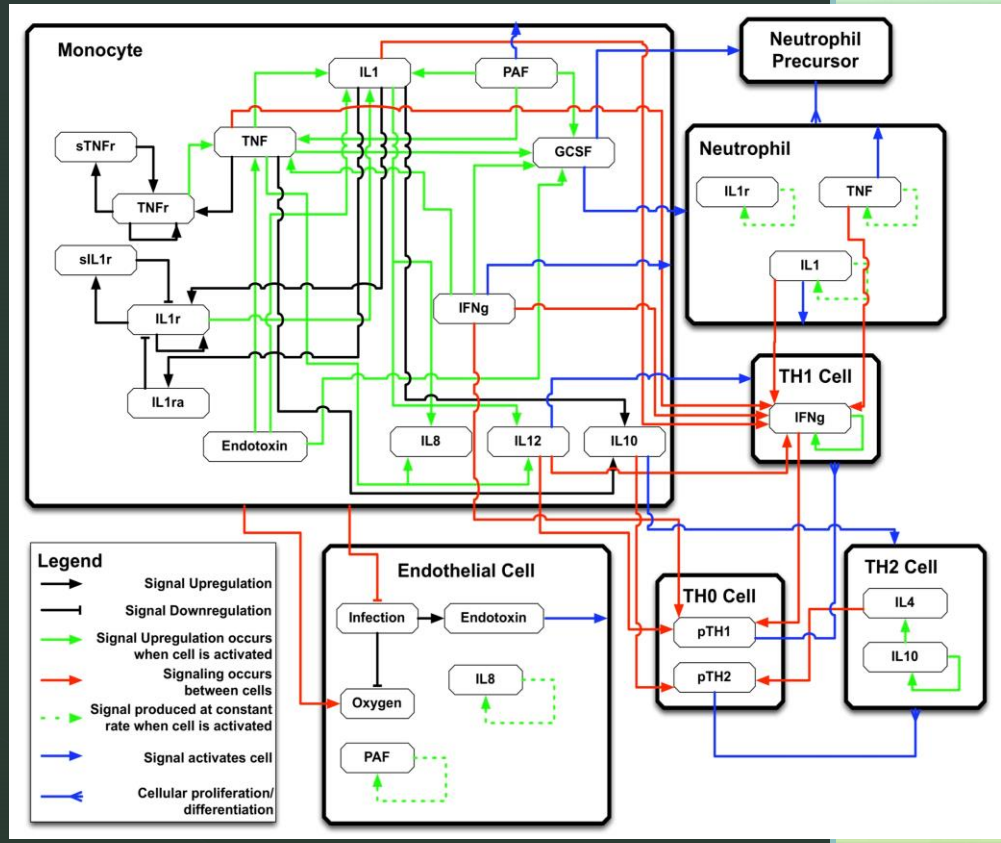
ORIGINAL RESEARCH article
Front. Physiol., 19 May 2021 | https://doi.org/10.3389/fphys.2021.662845

Utilizing the Heterogeneity of Clinical Data for Model Refinement and Rule Discovery Through the Application of Genetic Algorithms to Calibrate a High-Dimensional Agent-Based Model of Systemic Inflammation

Chase Cockrell* and Gary An
Departmen of Surgery, Larner College of Medicine, The University of Vermont, Burlington, VT, United States

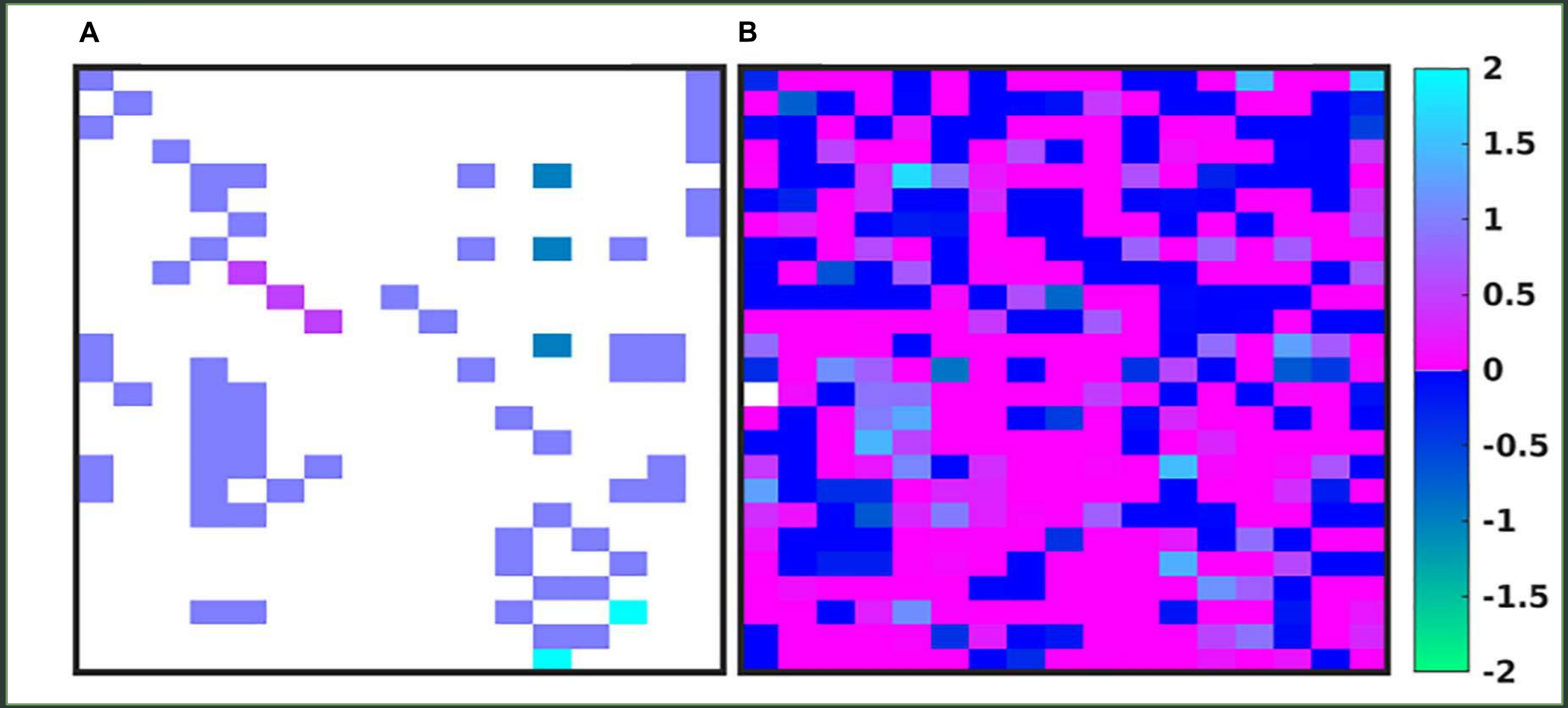
Example: Synthetic Multiplexed Molecular Time Series

- Simulation Model => Innate Immune Response ABM (CCM 2004)



- Data Source: USU/WRNMMC via SC2i TDAP protocol
  - 199 trauma patients: 92 developed ARDS; 107 controls wo ARDS
- Data:
  - Collected at t=0,1,3,7,14 days post-injury
  - Blood-serum cytokine profiles time series: IL-1b, IL-1ra, IL-6, IL-4, IL-8, IL-10, G-CSF, IFNg, and TNFa
  - Organ Failure Scores

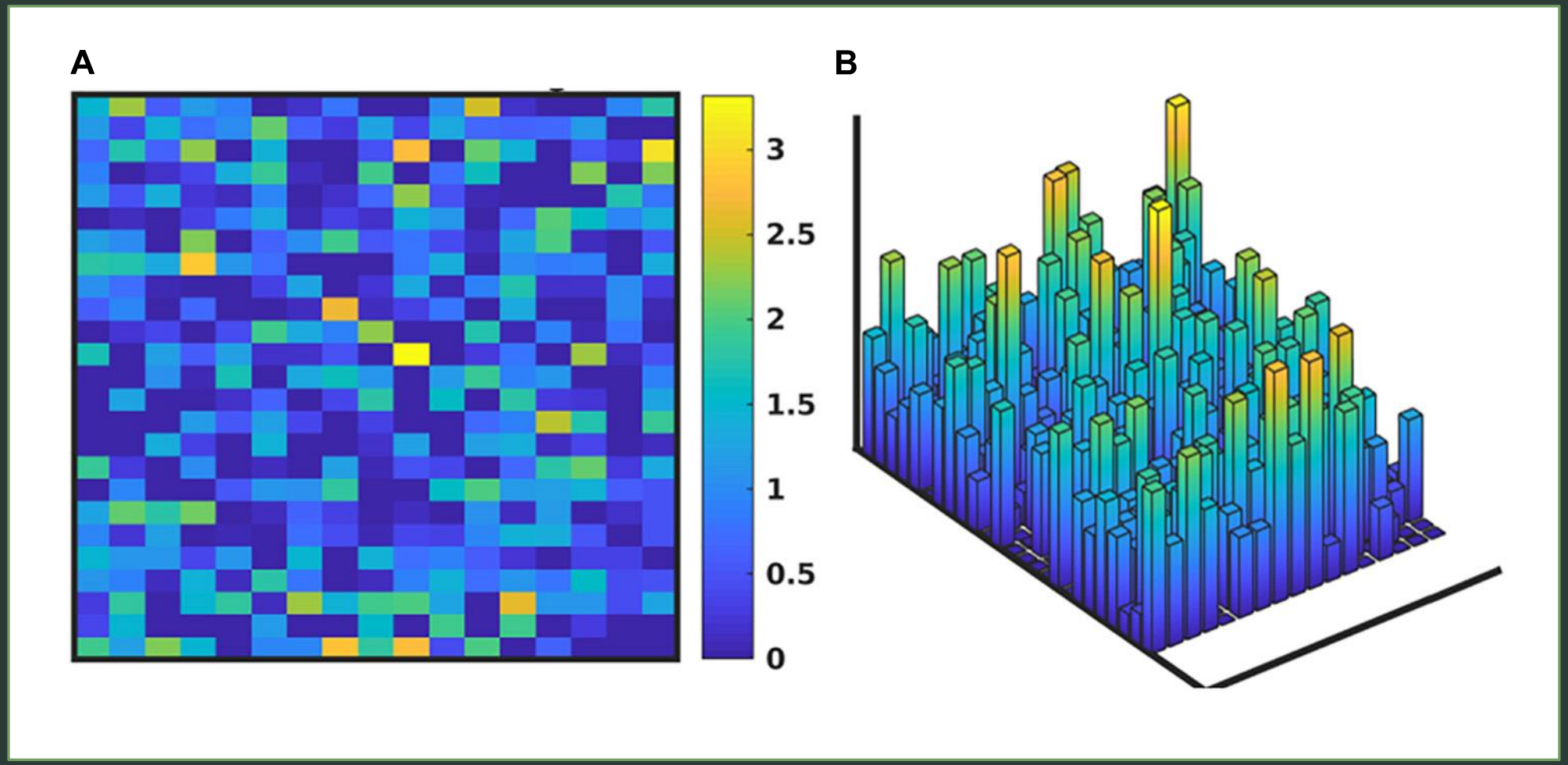
Results: Sample MRMs



Base IIRABM MRM

Calibrated IIRABM MRM

Results: Range of Ensemble MRMs

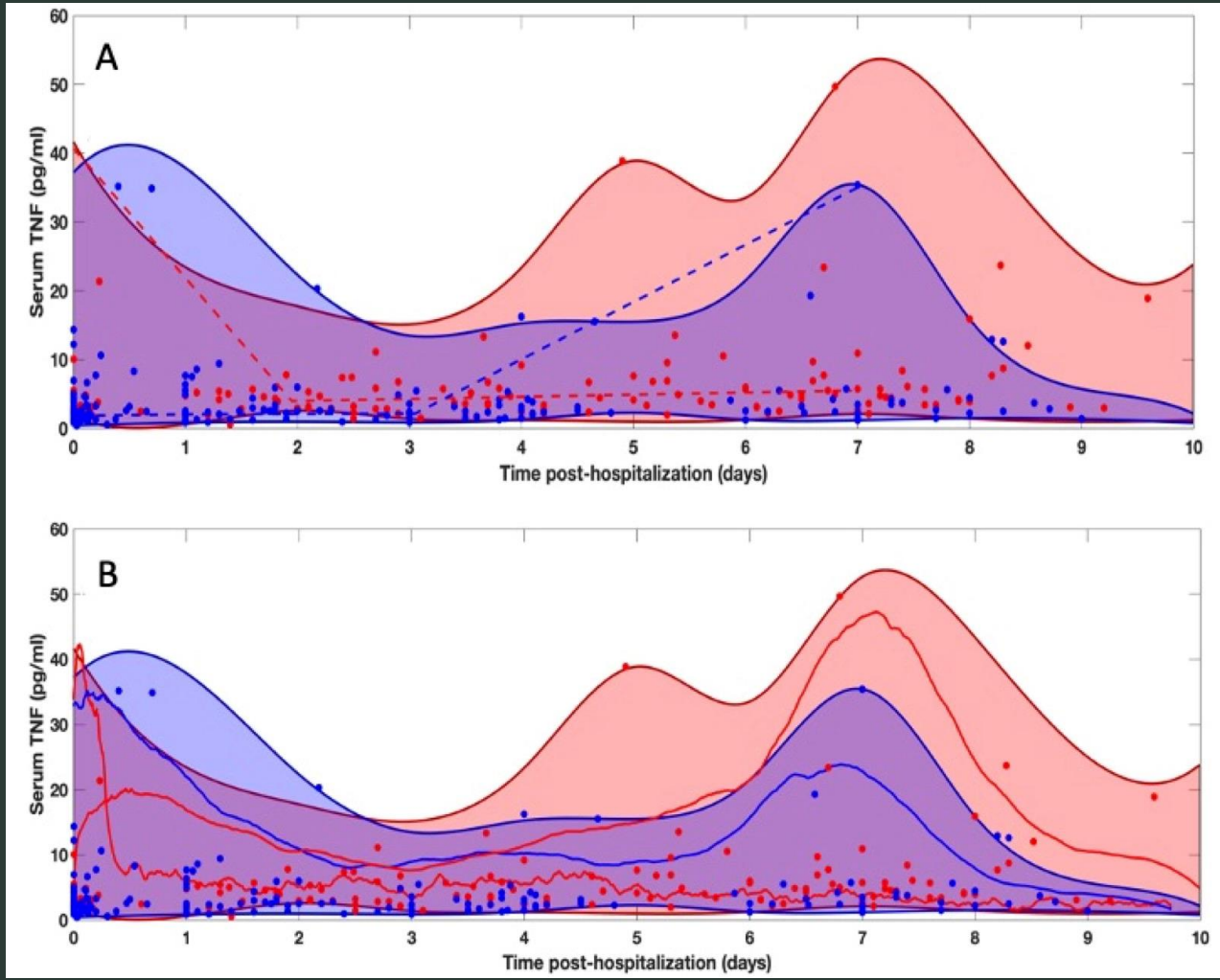


2d Heatmap of Value Ranges

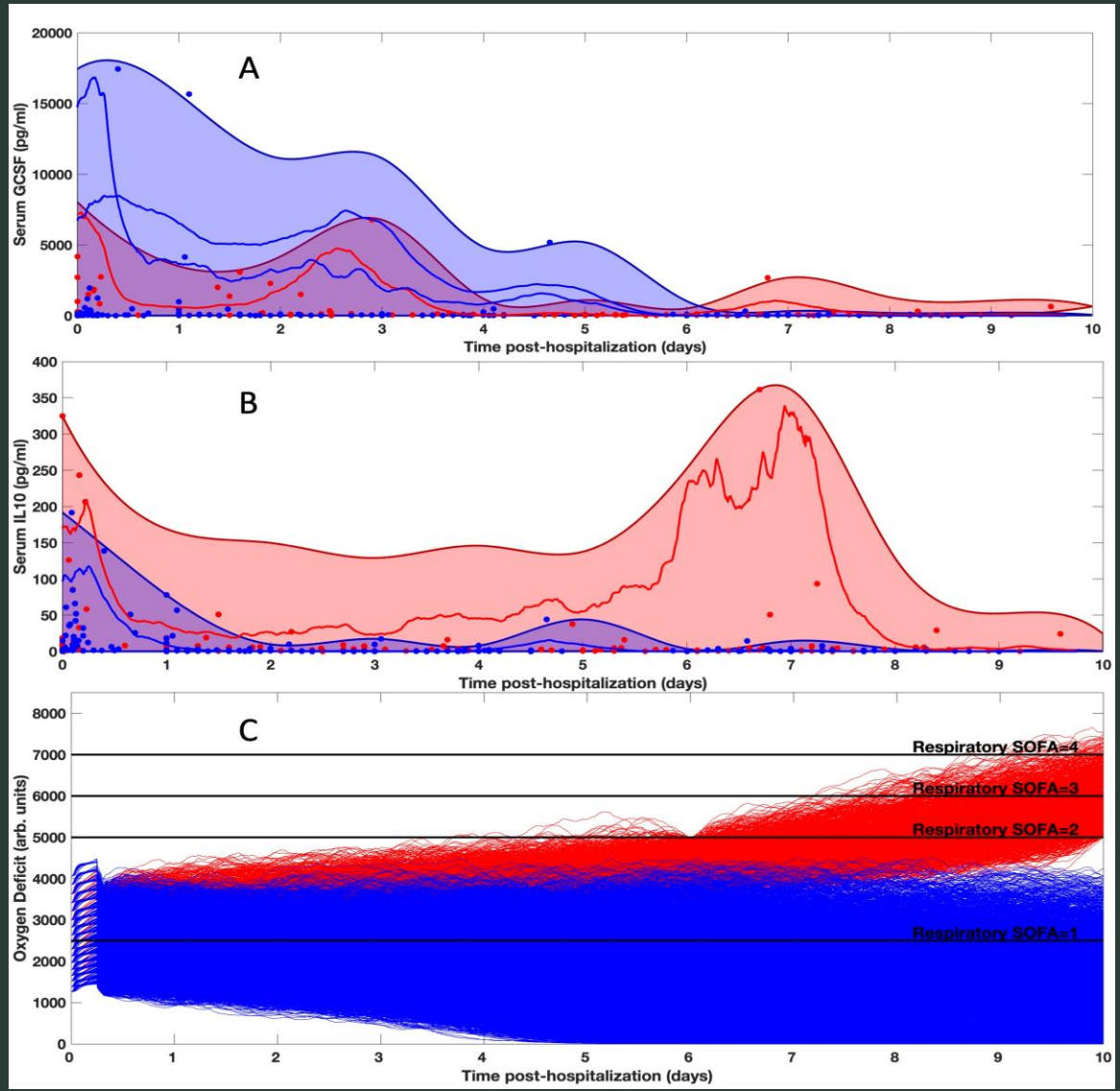
3d Depiction of Value Ranges

Results: Synthetic Trajectory Spaces

TNFa: Real and Synthetic Data



GCSF, IL-1 and SOFA Scores



Context of Control: Simulation-based Deep Reinforcement Learning (DRL)

- Simulation-based DRL = Game Playing AIs
- Hybrid systems => incorporate mechanism-based simulations to generate synthetic data
- Proposal: Simulation-based DRL for complex control discovery for biomedical problems (proof-of-concept for sepsis)

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

Brenden K. Petersen, Jiachen Yang, Will S. Grathwohl, Chase Cockrell, Claudio Santiago, Gary An, Daniel M. Faissol

Preparing for the next COVID: Deep Reinforcement Learning trained Artificial Intelligence discovery of multi-modal immunomodulatory control of systemic inflammation in the absence of effective anti-microbials

Dale Larie, Gary An, Chase Cockrell

doi: https://doi.org/10.1101/2022.02.17.480940

Context of Control: Simulation-based Deep Reinforcement Learning (DRL)

- Simulation-based DRL = Game Playing AIs
- Hybrid systems => incorporate mechanism-based simulations to generate synthetic data
- Proposal: Simulation-based DRL for complex control discovery for biomedical problems (proof-of-concept for sepsis)

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

Brenden K. Petersen, Jiachen Yang, Will S. Grathwohl, Chase Cockrell, Claudio Santiago, Gary An, Daniel M. Faissol

Preparing for the next COVID: Deep Reinforcement Learning trained Artificial Intelligence discovery of multi-modal immunomodulatory control of systemic inflammation in the absence of effective anti-microbials

Dale Larie, Gary An, Chase Cockrell

doi: https://doi.org/10.1101/2022.02.17.480940

Repurposed multi-modal control of sepsis wo effective antibiotics

- Validated Model of Sepsis => ”game” for DRL AI to win
  - “Novel Pathogen” => No effective anti-infective => 21 days in ICU
  - Hypothetical Sensing = Cytokine State q6 hrs
  - Hypothetical Interventions = ***Augment or inhibit*** q6 hrs
  - “Moves”: Existing approved anti-cytokine meds*
    - *TNF, IL1, IL2, IL4, IL8, IL12, IFNg
- Reward Functions
  - Step Reward = Reduce aggregate tissue damage
  - Terminal Reward = Successful Recovery at 21 days

Results

- 6000 Training Episodes for Convergence
- Baseline Mortality = **39%** (61% Recovered) => Controlled Mortality = **10%** (90% Recovered)
- Required manipulating 6 mediators at variable levels at variable times (TNF, IL1, IL4, IL8, IL12, IFNg)
- Robustness of Policy => Tested across regions of parameter space

Parameterization	Uncontrolled Recovery Rate	Controlled Recovery Rate	Improvement
Test 1: (0.1,1,3,20)	25%	81%	56%
Test 2: (0.12,1,1,32)	16%	56%	40%
Test 3: (0.08,2,1,23)	19%	52%	33%
Test 4: (0.12,2,1,28)	37%	83%	46%

The baseline and controlled mortality rates (MR) for IIRABM parameterizations upon which the DRL algorithm was *not* trained are presented here. Parameterizations are defined as (host resilience, microbial invasiveness, microbial toxigenesis and initial injury size).

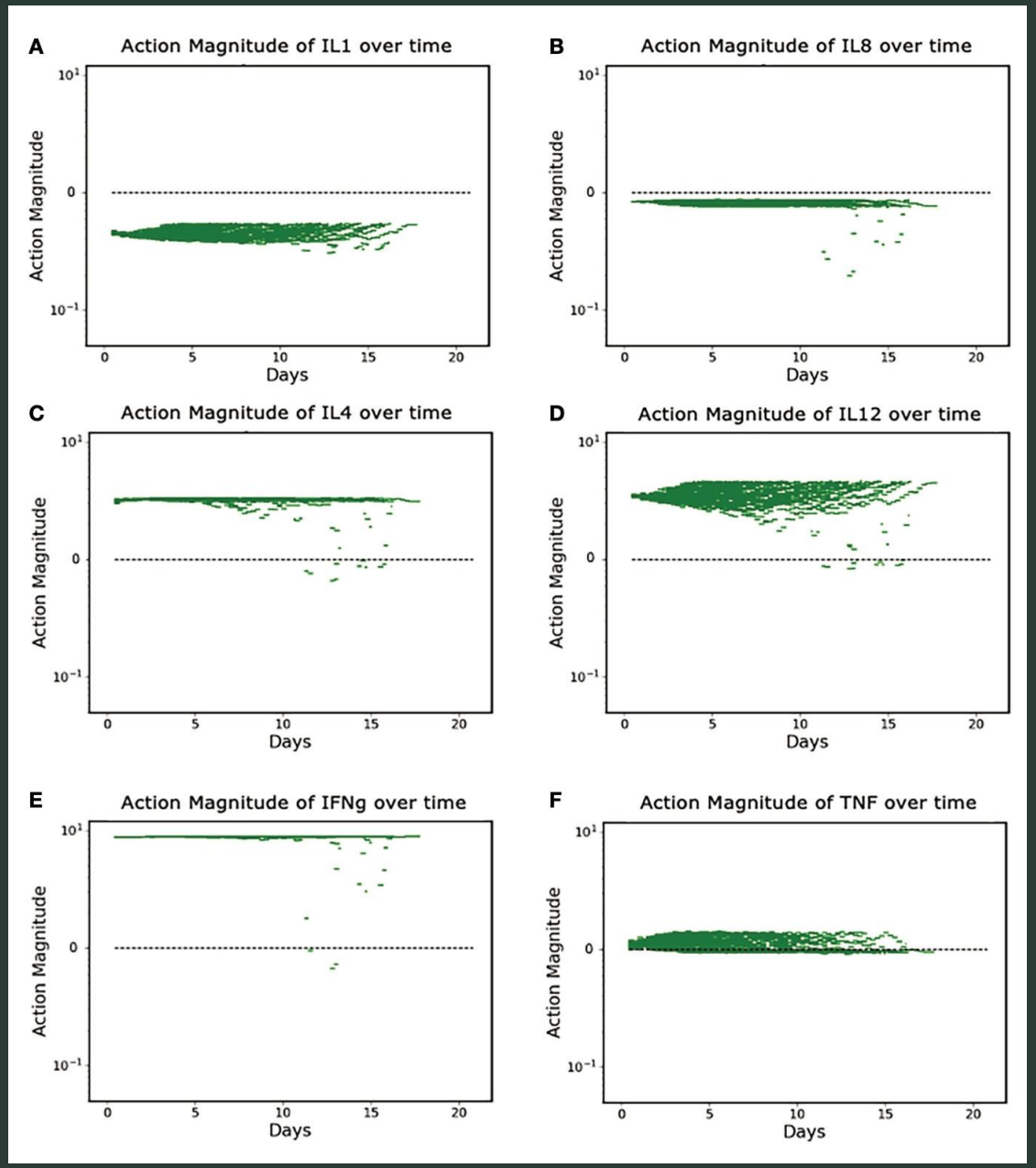
Results

- 6000 Training Episodes for Convergence
- Baseline Mortality = **39%** (61% Recovered) => Controlled Mortality = **10%** (90% Recovered)
- Required manipulating 6 mediators at variable levels at variable times (TNF, IL1, IL4, IL8, IL12, IFNg)
- Robustness of Policy => Tested across regions of parameter space

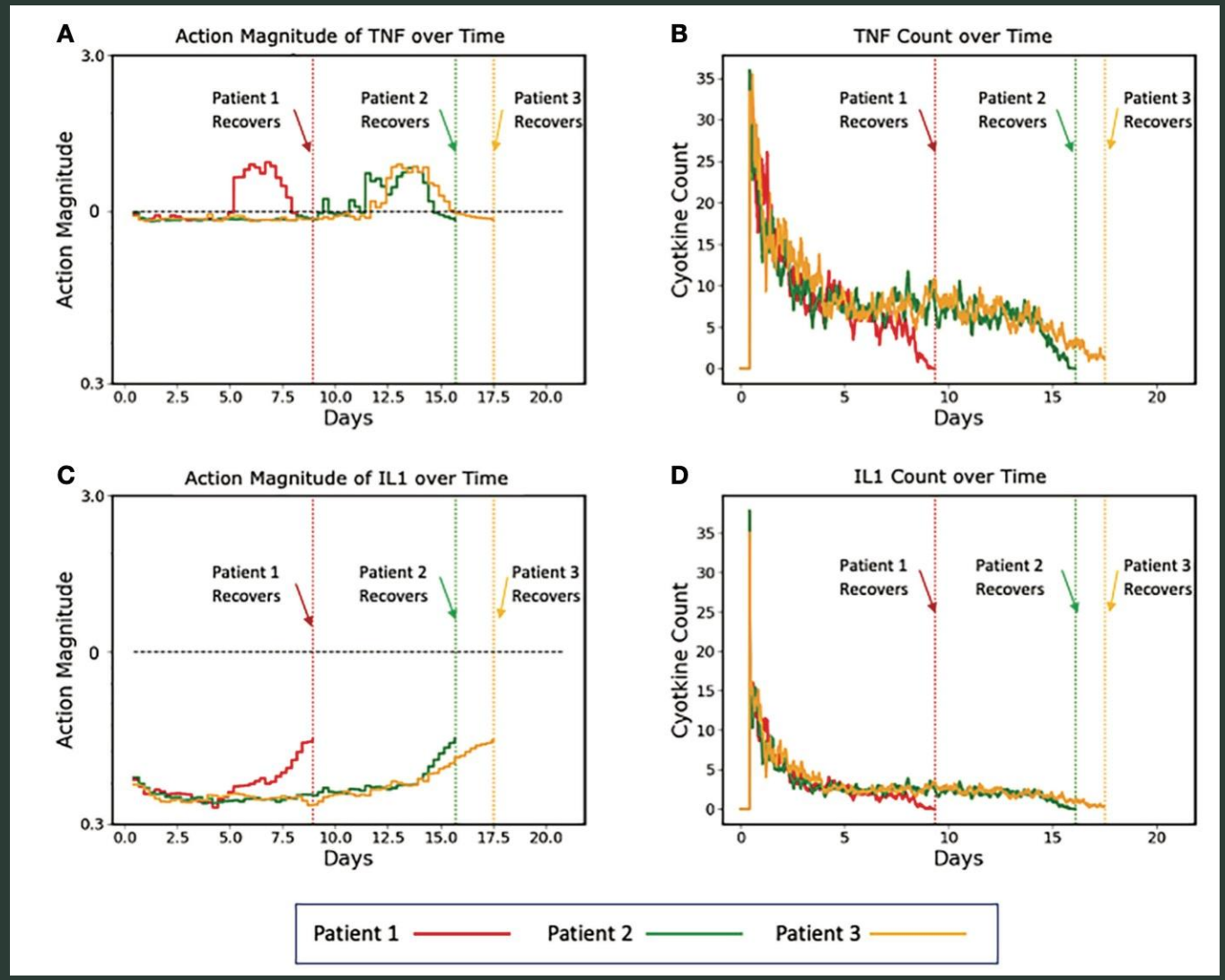
Parameterization	Uncontrolled Recovery Rate	Controlled Recovery Rate	Improvement
Test 1: (0.1,1,3,20)	25%	81%	56%
Test 2: (0.12,1,1,32)	16%	56%	40%
Test 3: (0.08,2,1,23)	19%	52%	33%
Test 4: (0.12,2,1,28)	37%	83%	46%

The baseline and controlled mortality rates (MR) for IIRABM parameterizations upon which the DRL algorithm was *not* trained are presented here. Parameterizations are defined as (host resilience, microbial invasiveness, microbial toxigenesis and initial injury size).

Control Actions over Time



Variable Patient Trajectories

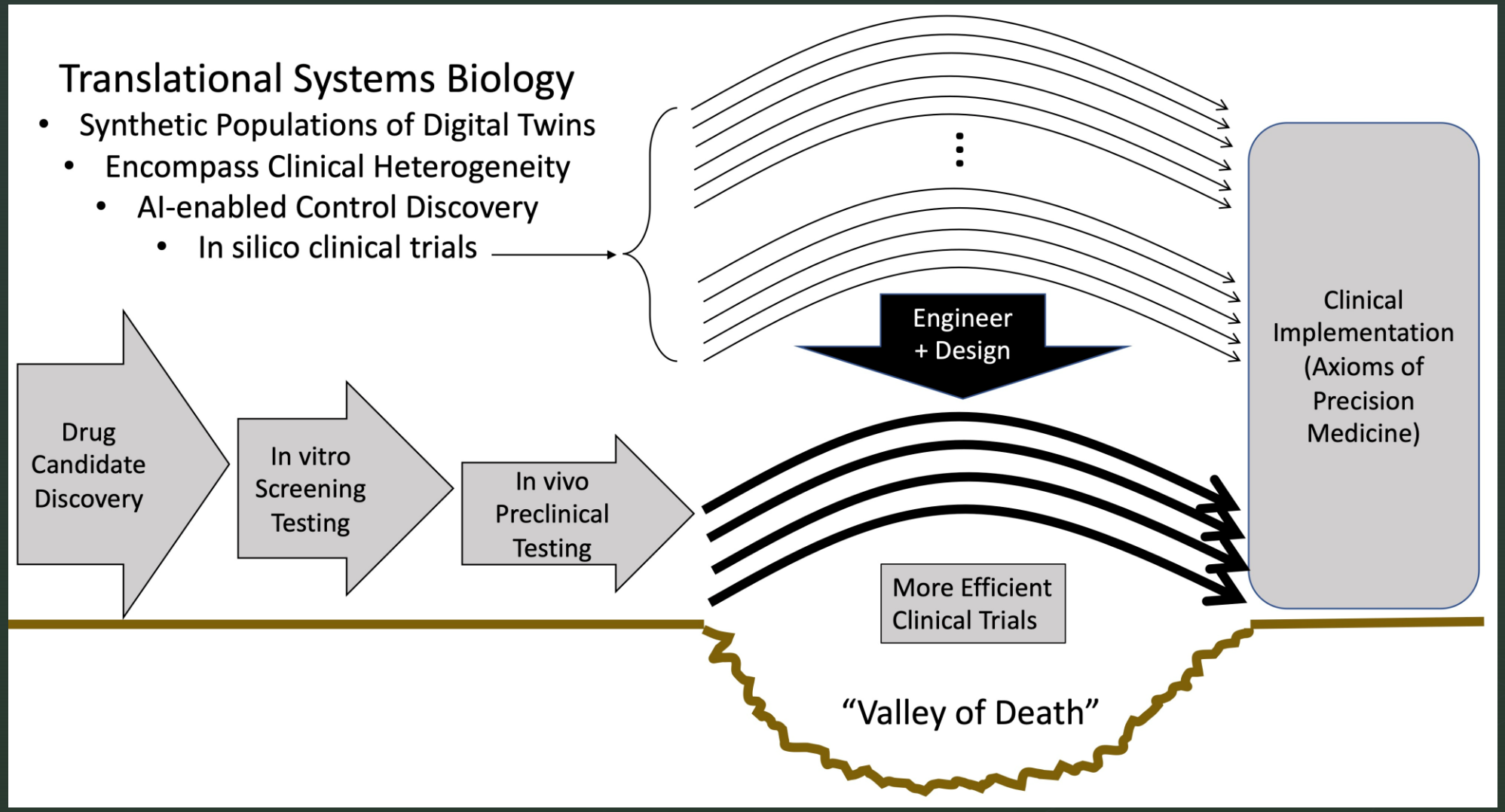


Conclusions

- Crossing Valley of Death: Upstream of QSP
- Rare diseases and “True” Precision Medicine => Share approaches?
  - Need to understand how the system works (mechanisms)
  - Need to reconcile individual/biological heterogeneity in the face of epistemic uncertainty
  - Extended Maximal Entropy Principle: Extend to characterizing unknowns adjacent to represented knowledge
  - *Be comfortable operating over very large (but not infinite!) parameter spaces*
  - ***Find Robust control strategies (e.g. clinically effective drugs) => take advantage of biological structure***

Bridging The Valley of Death

- Augmented Throughput over the Valley of Death



Finis