A Machine Learning Based Approach for Toxicity Predictions in Immuno-Oncology

Kamrine Poels, PhD

Senior Scientist in ECD-Quantitative Systems Pharmacology, Pfizer

September 15th, 2022

Expanded modeling and simulation approaches in support of Model-Informed Drug Development

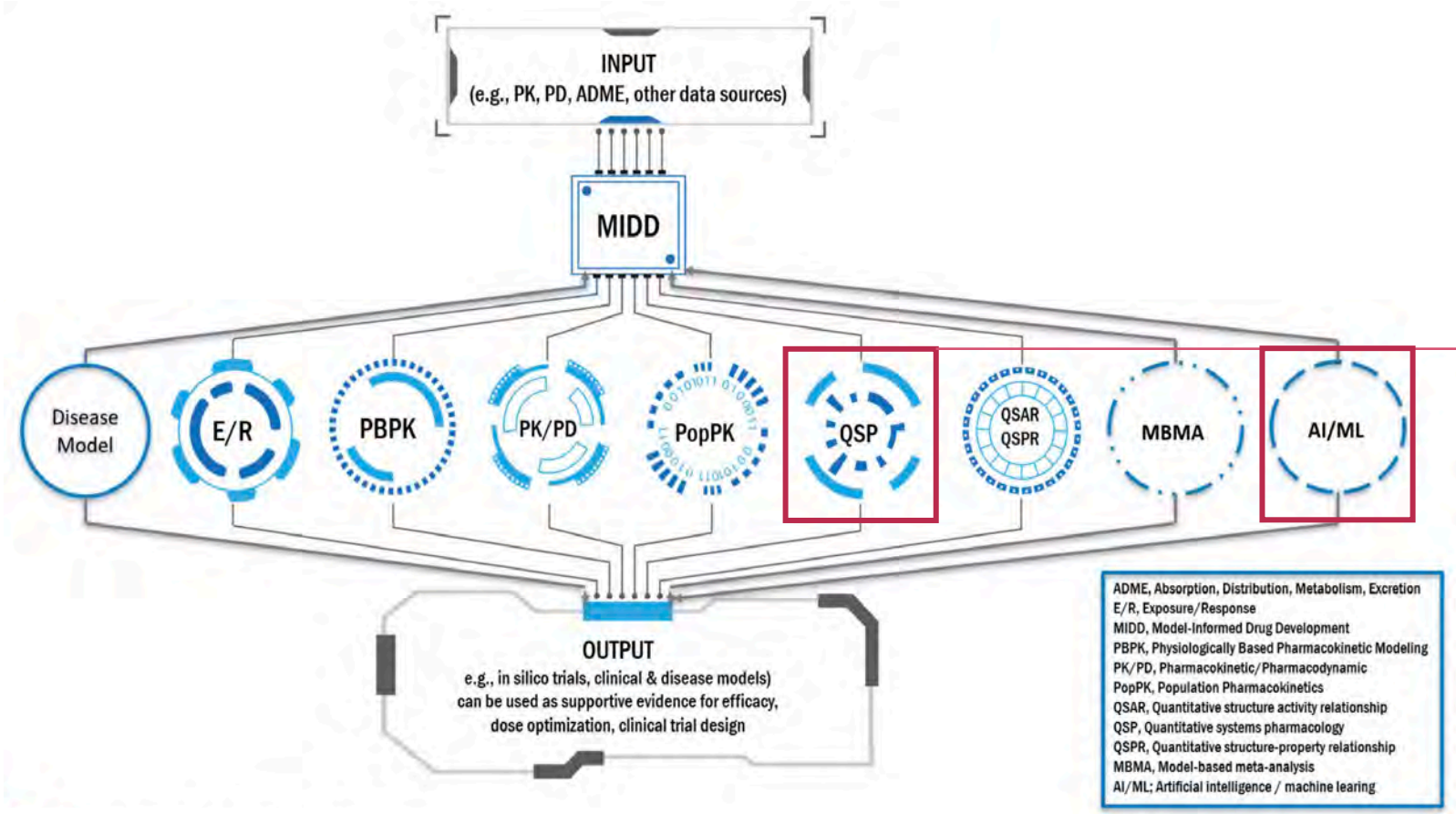


Figure credit: Kimberly Bergman, Hao Zhu (FDA)

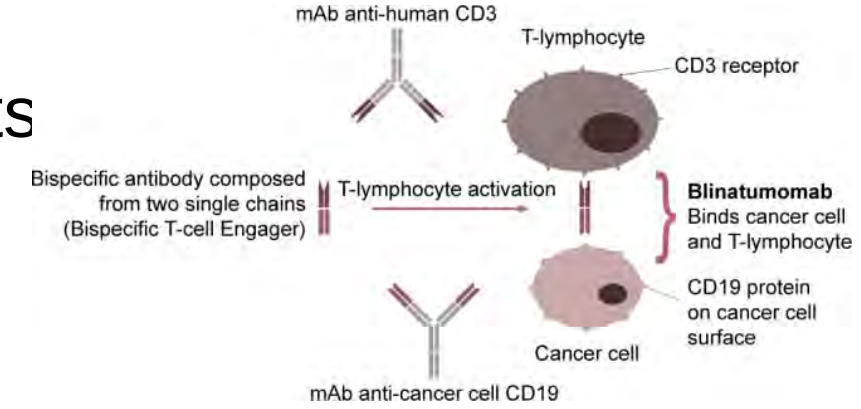
Best of both worlds QSP + ML/AI:

"The CDER investigators suggest that when scientists from the machine learning and mechanistic modeling disciplines interact and learn from each other, they will find many overlapping concepts, equations, and ways to visualize model structure architectures, and that machine learning models and more traditional mechanistic ones may complement each other to improve the efficiency of drug development and optimize treatments for individual patients."

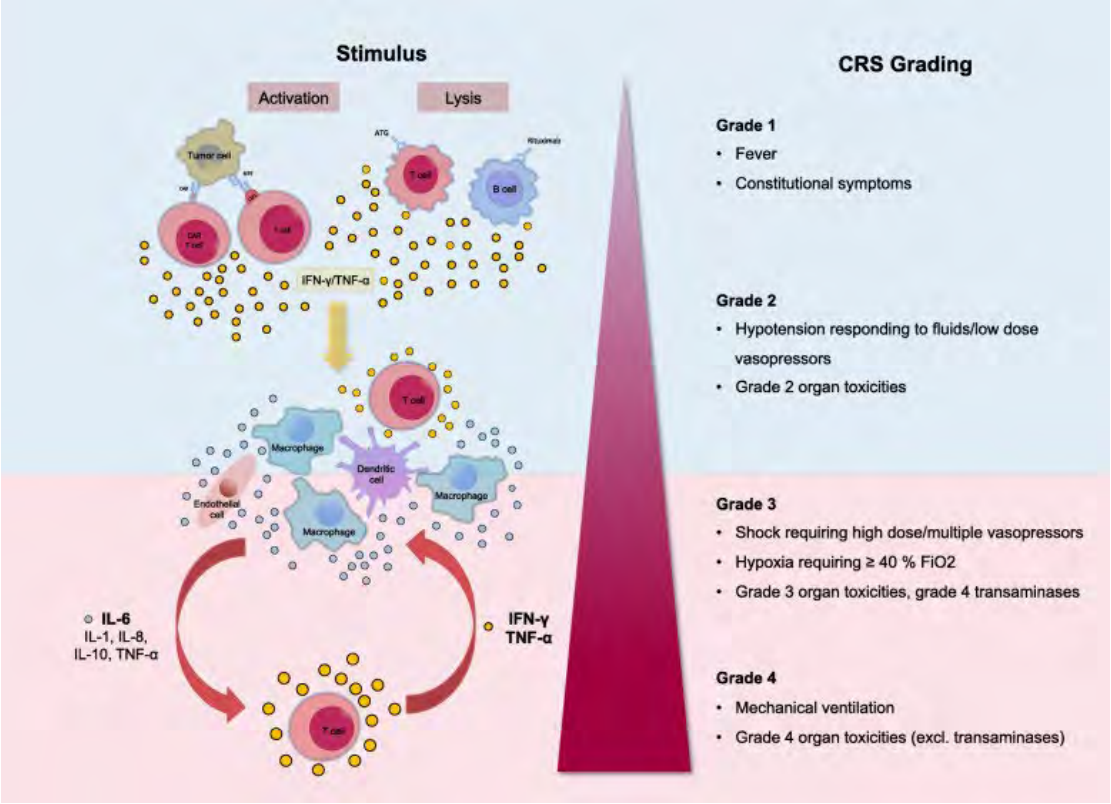
Source: https://www.fda.gov/drugs/regulatory-science-research-and-education/new-approach-pharmacometrics-recurrent-neural-networks-modeling-drug-exposure-and-drug-response

Cytokine release syndrome can occur during bispecific antibody treatment for cancer patients

Bispecific antibody treatment is a promising cancer treatment



- Bispecific antibody (bsAb) links a T cell and a cancer cell
  - Binds regardless of T cell (binds portion of TCR)
  - Binding proximity triggers T cells, release tumor killing agents
- T cell activation can cause release of toxic cytokines
  - On-target activation leads to non-target activation and excessive cytokine release
  - Excessive cytokine release leads to CRS → graded response



[1]Sedykh et al. Drug Des Devel Ther. (2018) 12, 195-208.
[2]Shimabukuro-Vornhagen et al. J Immunother Cancer (2018) 6, 56.

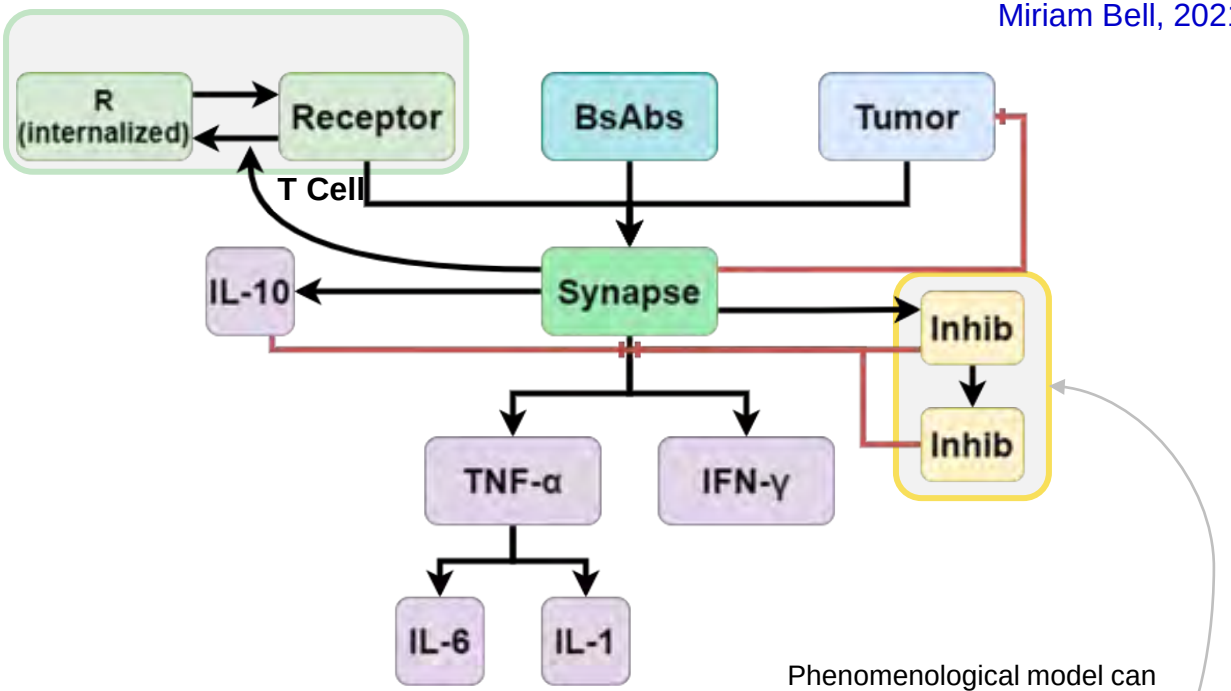
QSP model of CRS with priming: from bench to clinic

Miriam Bell, 2021

Key biological features informed from *in vitro* experiments

- Independent thresholds for cytotoxicity and toxic cytokine release
  - Tumor killing
  - Cytokine release
- Cytokine feedback
  - IL-10 inhibits the production of other cytokines
- Internal inhibition mechanism for T cell desensitization
- Dose priming can reduce CRS but must maintain tumor killing efficacy[1]
  - Mechanism of dose priming is incompletely understood
    - Literature[3] suggests it's a combination of TCR downregulation and a downstream switch

Model can capture key features for drug efficacy and cytokine release

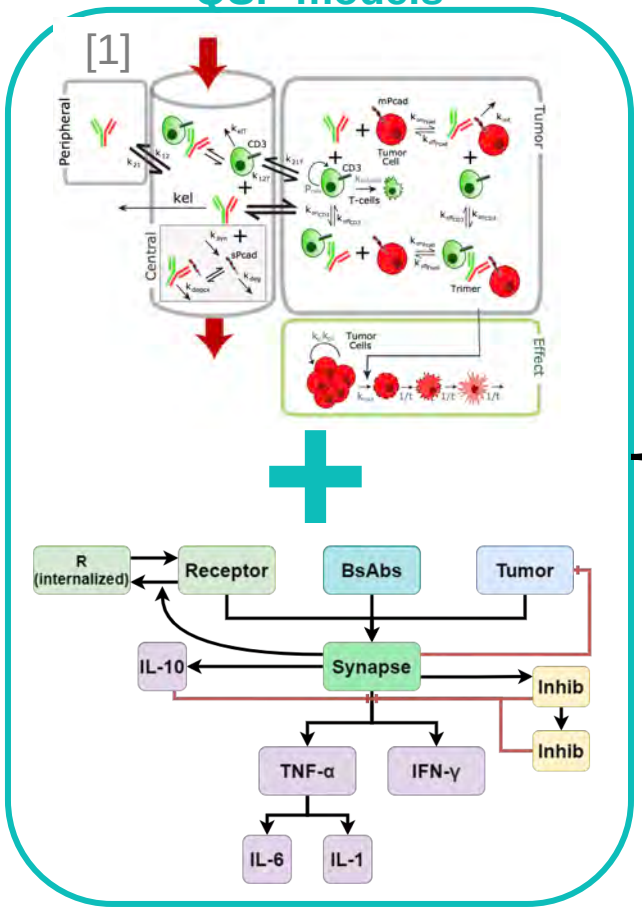


Phenomenological model can capture dose priming dynamics but need biological basis – "black box" for desensitization

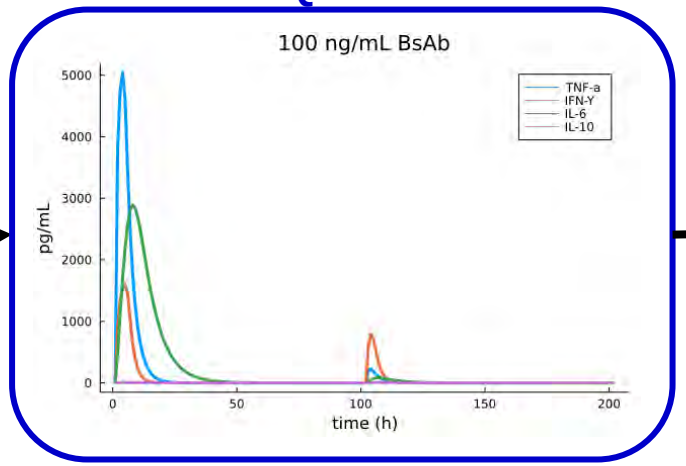
[1]Chen et al. Clin Transl Sci (2019) 12, 600-608.
[2]Li et al. Sci transl Med (2019) 11 (508)
[3]Trendel et al. Sci Signaling (2021) 14 (666),

Hybrid QSP-AI modeling for bsAb therapy: ML for CRS classification

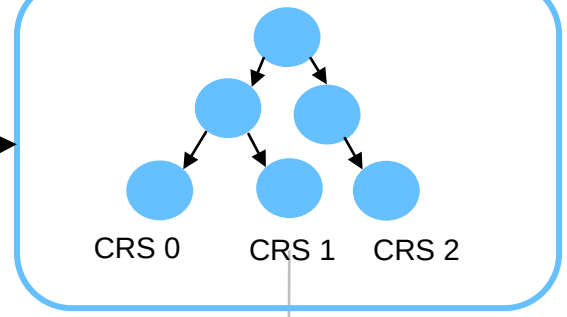
Clinical efficacy and CR QSP models



Cytokine dynamics from QSP Model



CRS Classifier



Efficacy simulations via Virtual Populations[2]

Recommendations based on dose optimization

[1]Betts et al. *The AAPS Journal* (2019)
[2]Allen et al. *CPT Pharmacometrics Syst. Pharmacology* (2016)

ML methods for analysis of clinical data for cytokine to grade mapping

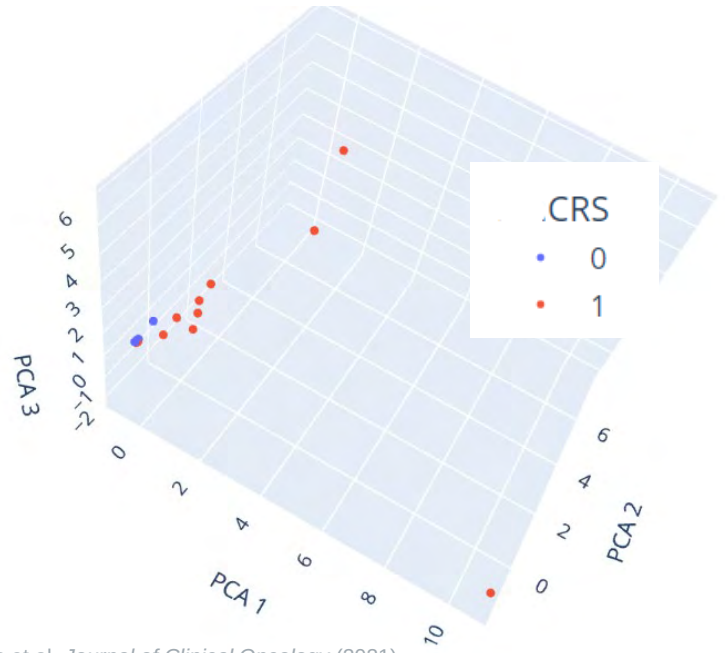
Proof of concept: cytokine signaling may be sufficient for CRS prediction

Emily Nieves, 2022

PCA separates patients with CRS from non-CRS

- Optimal featurization: rate of change of cytokine from baseline to peak

IV Dose Escalation Cohort CRS vs. Non-CRS



Bahlis et al. *Journal of Clinical Oncology* (2021)

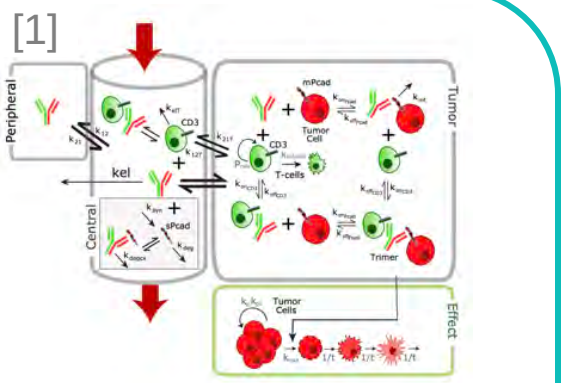
Variable importance

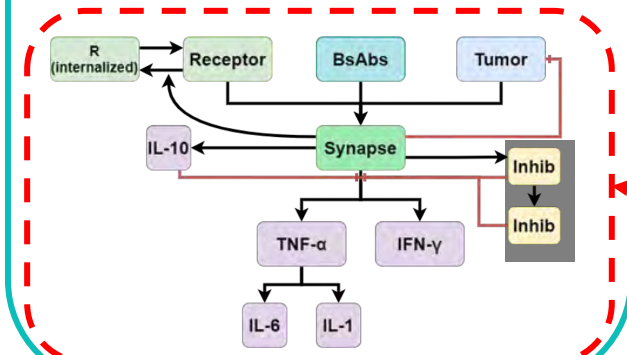
- **Random forest algorithm** to classify CRS grade
  - Used rate of change of each cytokine from baseline to peak as predictor
  - Adjusted for dosing
  - Identified **IL-6** and **IL-10** as key predictors
- Algorithm correctly classifies all non-CRS test cases, but mislabels CRS grades 1 and 2
  - Average cross-validated accuracy: 60%
  - Average cross-validated AUC: 0.80

Hybrid QSP-AI modeling in bsAb therapy: uncover the biology

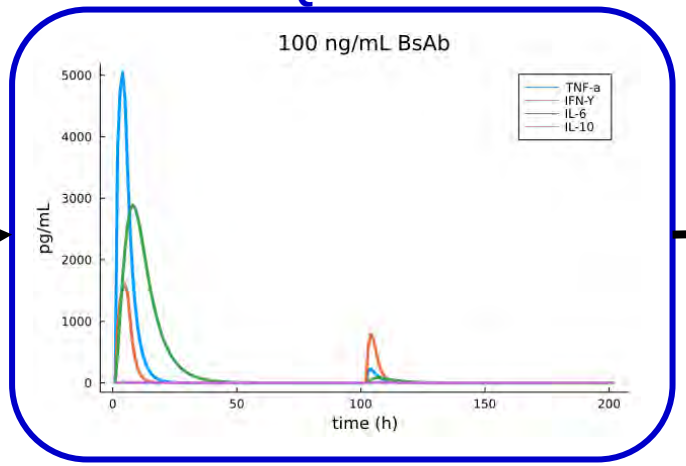
Clinical efficacy and CR QSP models

[1]

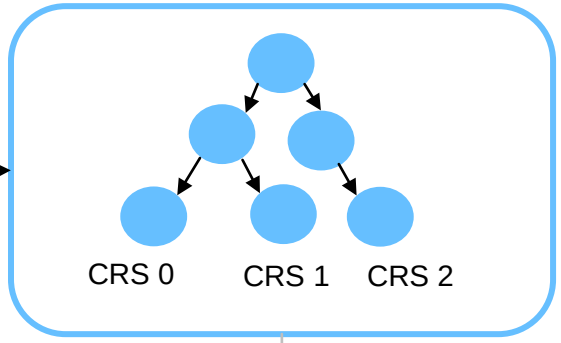




Cytokine dynamics from QSP Model



CRS Classifier



Efficacy simulations via Virtual Populations[2]

Recommendations based on dose optimization

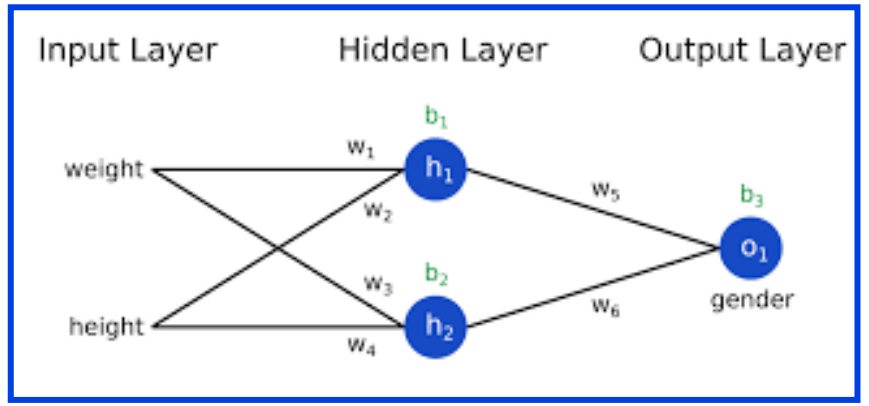
Use more novel AI/ML methods to discover structure of mechanisms?

[1]Betts et al. *The AAPS Journal* (2019)
[2]Allen et al. *CPT Pharmacometrics Syst. Pharmacology* (2016)

Automated model discovery: using data to learn missing terms of a system of ODEs

- **Neural Networks (NN)**
  - Inputs are weighted and summed, activation functions are applied to capture non-linear relationships
  - Capture complex patterns from data
  - With enough layers and units, NNs are universal function approximators

NN Example



- **Methodology**
  - Replace parts of differential equations with neural nets and optimize neural nets so output of ODE model matches data[1]
  - Use sparse regression to recover the equations of the additional term needed to reproduce data
  - With Julia's SciML ecosystem we can fit the parameters of the NN within an ODE problem

$$dx/dt = k_1 y + NN$$

$$\downarrow$$

$$NN \approx -k_2 x$$

[1]Rackauckas et al. 2020 preprint

What can automated model discovery do to support mechanistic modeling?

Example: Term related to quarantine strength recovered from COVID-19 data using automated model discovery[1]

- Adapted SIR epidemiological model to include effects of quarantine
- Use automated model discovery to learn a function for the strength of quarantine (Q(t)) from CDC data for different US states
  - Q(t) is a composite of testing rate and quarantine policies
- Incorporate incompletely understood and complex mechanisms into models with data without the opacity of pure ML methods

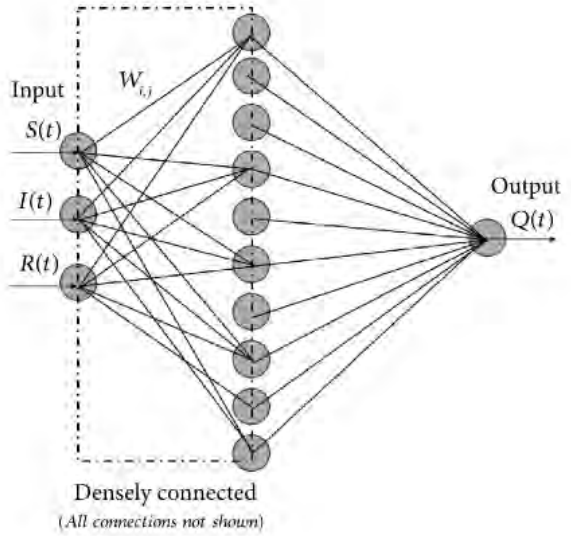
$$\frac{dS}{dt} = -\frac{\beta S(t) I(t)}{N},$$

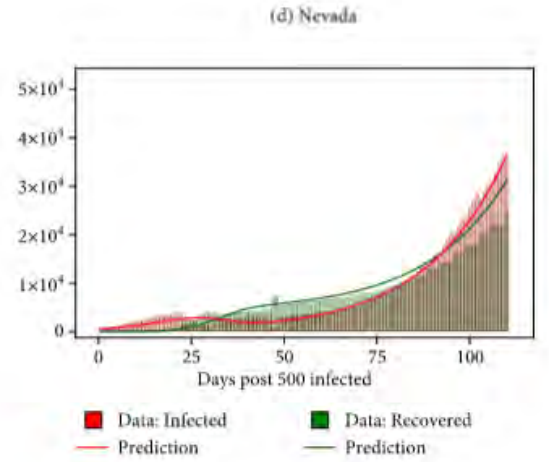
$$\frac{dI}{dt} = \frac{\beta S(t)\ I(t)}{N} - (\gamma + Q(t)) I(t)$$

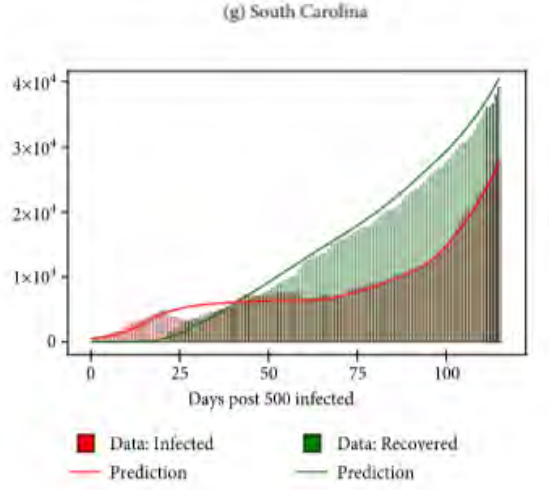
$$= \frac{\beta S(t) I(t)}{N} - (\gamma + \mathrm{NN}(W,U)) I(t),$$

$$\frac{dR}{dt} = \gamma I(t) + \delta T(t),$$

$$\frac{dT}{dt} = Q(t)\ I(t) - \delta T(t) = \mathrm{NN}(W,U)\ I(t) - \delta T(t).$$







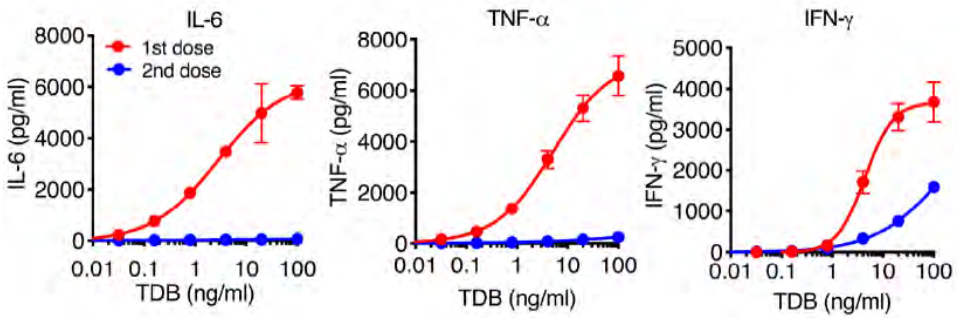
[8] Dandekar et al., *Health Data Sci* 2021

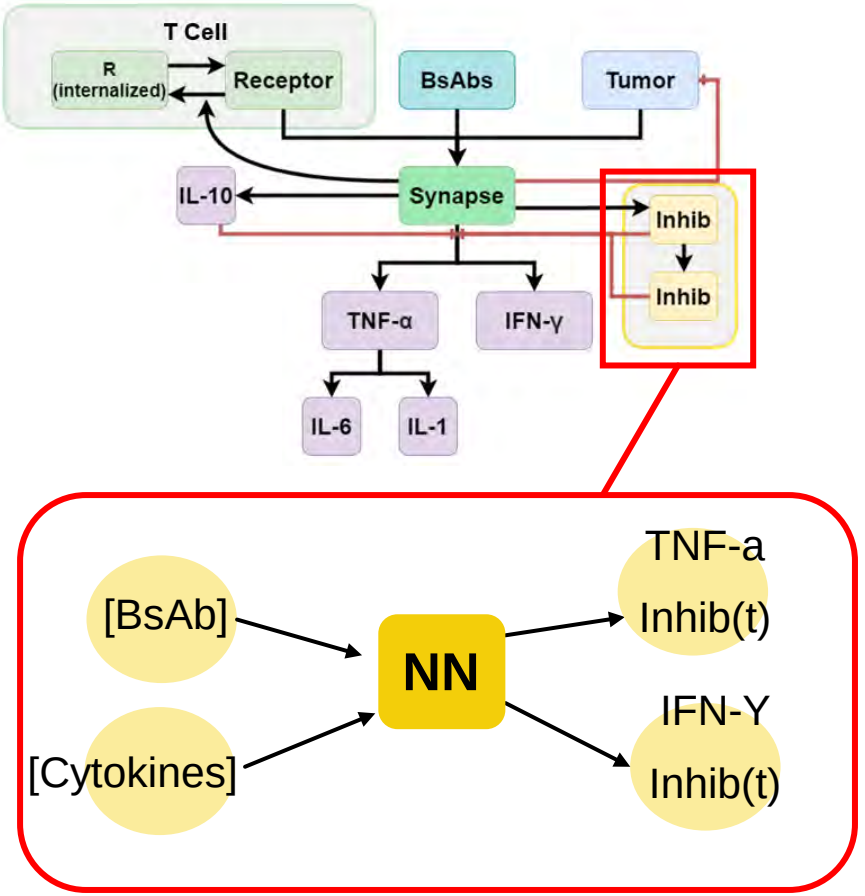
[1]Dandekar et al., *Health Data Sci* 2021

Proposal: Model discovery to learn dose priming inhibition term

On-going work with PhD candidate Emily Nieves and Chris Rackauckas, PhD

- Apply automated model discovery to add mechanistic detail to dose priming component and improve fits
  - Learn inhibition term that varies with time from data
  - Inhibition term will likely be a composite of TCR downregulation and downstream mechanisms related to desensitization
  - Sparse regression to recover equations from NN





Li et al. *Sci Trans Med* 2019

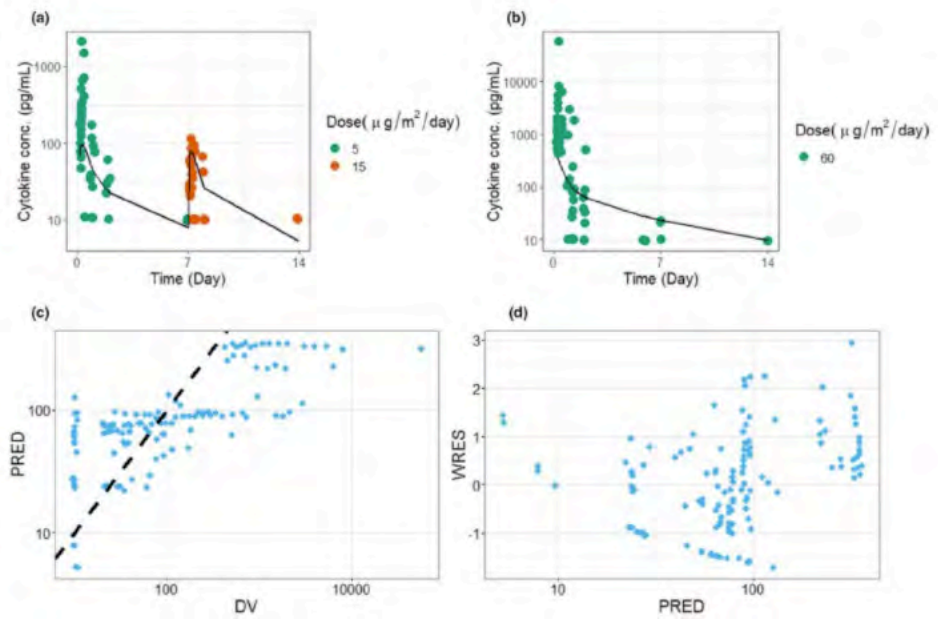
Acknowledgements

- **Pfizer**
  - *Blerta Shtylla*
  - Claire Zhao, Subha Madhavan
  - Rohit Rao, Richard Allen
  - C.J. Musante, Gianluca Nucci, Sandeep Menon
- **MIT**
  - *Emily Nieves*, Chris Rackauckas
- **UCSD**
  - *Miriam Bell*

Supplementary slides

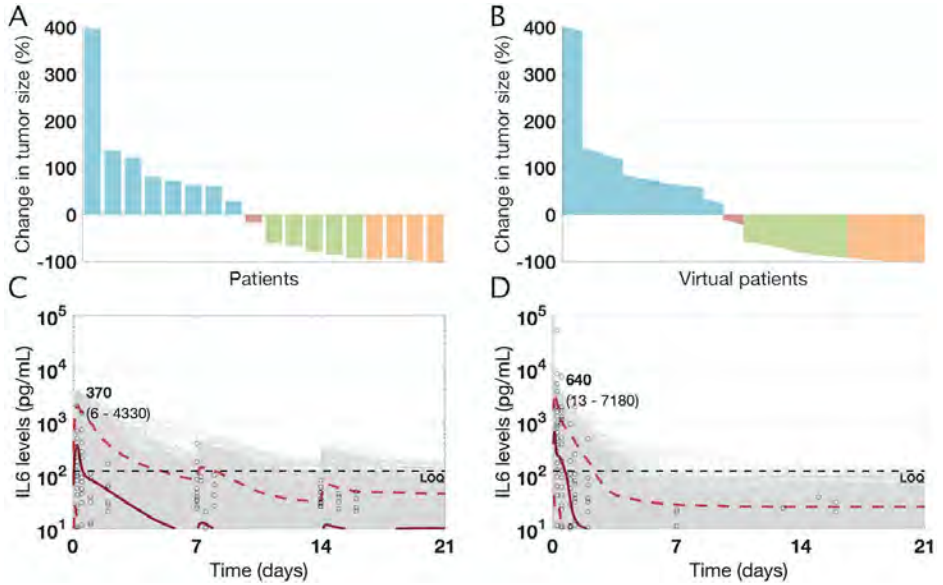
Current mathematical models of dose priming insufficient for predictive framework

Chen et al.



- Recapitulates observed cytokine release upon dose priming
- Phenomenological (lacking mechanistic detail)
  - Cannot capture temporal dynamics: cytokine effects artificially reset after each dose
- Unable to model interplay between safety and efficacy signals

Hosseini et al.



- Able to predict both safety and efficacy signals due to dose priming
- Dependent only on decreasing tumor burden for dose priming effect - Li et al showed priming independent of tumor burden in vitro

1 Chen et al. Clin Transl Sci (2019) 12, 600-608.
2 Hosseini et al. NPJ Sys Biol Appl 2020 6 (1)