An Open-Source[1] Probabilistic Modeling Framework for Personal Pandemic Risk Assessment

Pandemonium

Authors: Richard Hamlin, Nancy Ramirez-Herrera, Yunge Li, Glenn Tang, Pablo Paiewonsky, Jiaying Li, Parnika Lal, Aakash Upadhyay. Quantum Risk Analytics, Inc.

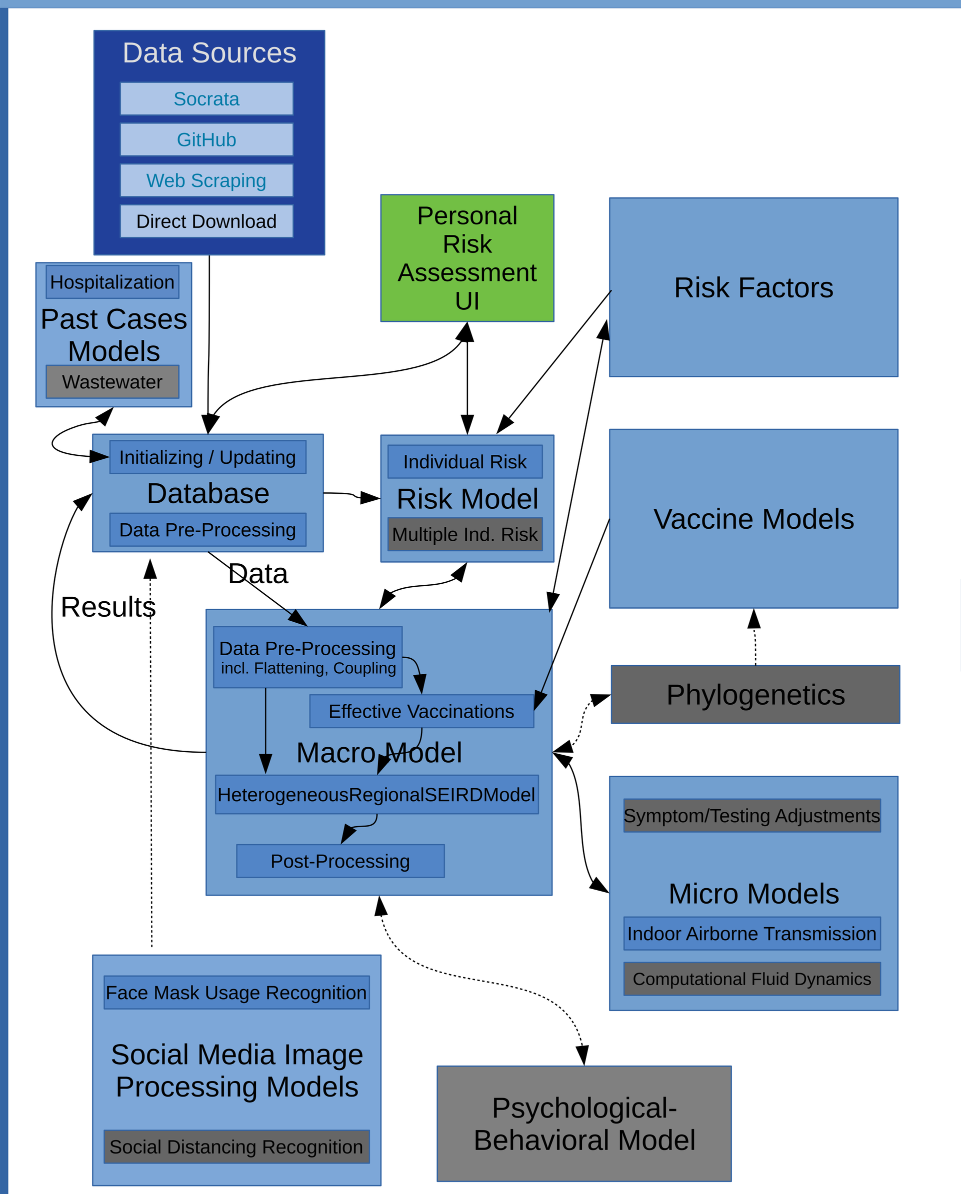
Abstract

Pandemonium is a machine learning-based framework that incorporates various epidemiological models and data to assess respiratory infections disease risks. The currently-used models and data are applied toward a web-based app for individual risk evaluation of **COVID-19** outcomes, to help users optimize **strategies for mitigating exposure.**

APP Description

Users will be able to estimate the risk of infections by entering COVID-related factors such as age, sex, chronic condition, and vaccine and location histories.

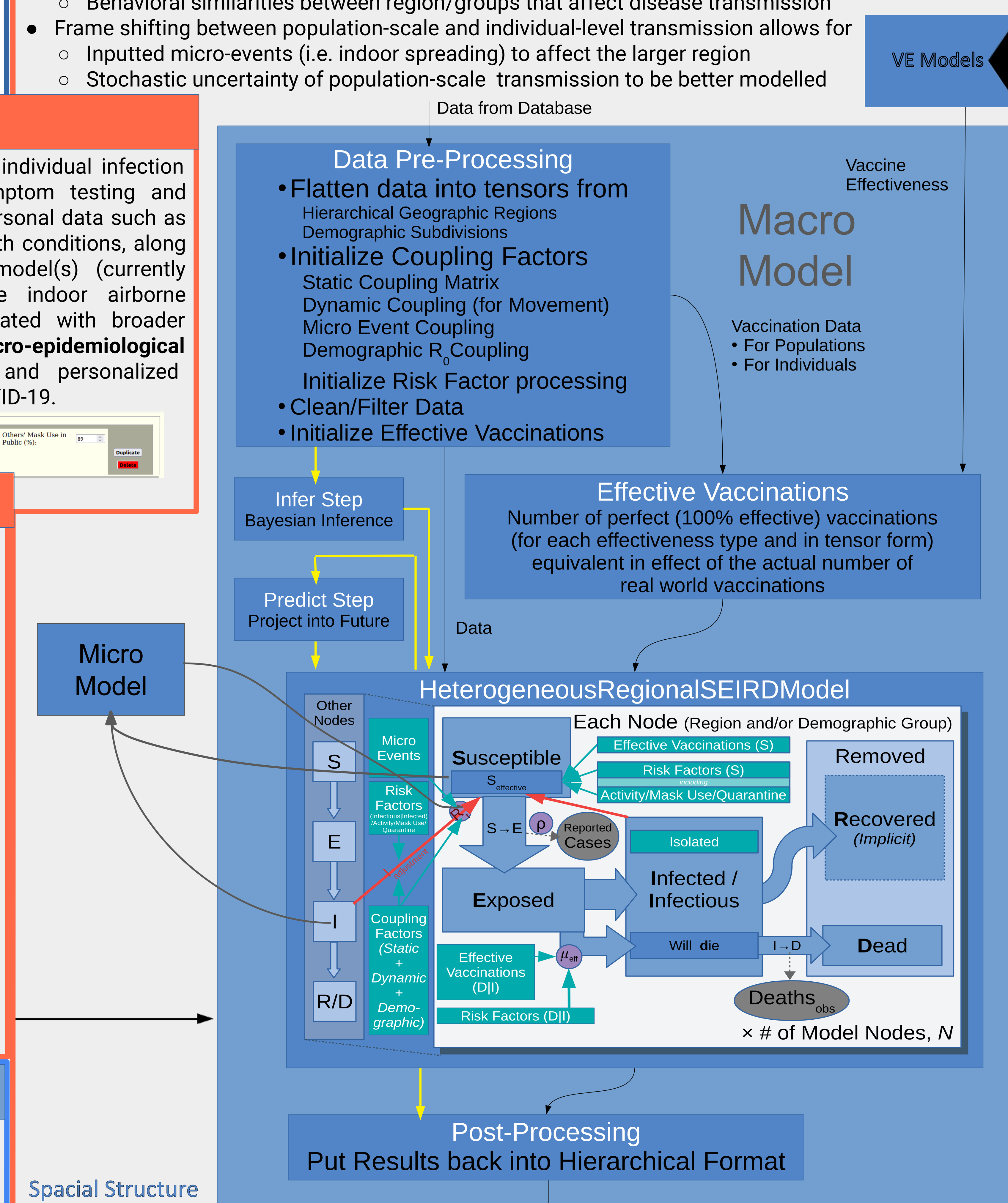
Epidemiological Model & Framework



- **Pandemonium's** framework combines several components, including:
  - **Vaccine Effectiveness (VE) model**
    - evaluates how vaccines impact infection risk
    - Currently using our own original **probabilistic VE model**.
  - **Risk factor framework**
    - Risk factor organization, aggregation, conversion.
  - **Macro model**
    - Main model
    - Simulator of disease dynamics

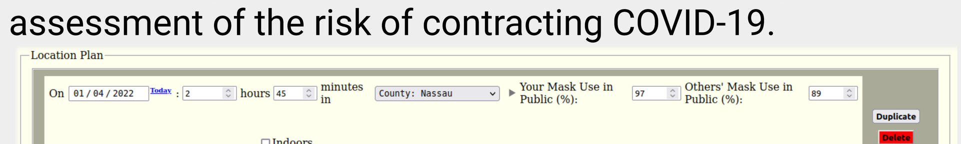
Macro Model

- Uses probabilistic programming to simulate population-level disease dynamics using **SEIRD** models, informed by risk factors and public health data.
  - **SEIRD** model parameters and intercompartmental flows can incorporate regional or group-specific risk factors such as age, vaccinations data, and prevalences of chronic health conditions.
- Models hierarchical geographic regions and demographic groups, including individuals.
- Uses coupling between region/groups via "coupling factors" to help simulate the effects from:
  - Flow of people between region/groups
  - The special movement of people between regions through time-varying "dynamic" coupling
  - Behavioral similarities between region/groups that affect disease transmission
- Frame shifting between population-scale and individual-level transmission allows for
  - Inputted micro-events (i.e. indoor spreading) to affect the larger region
  - Stochastic uncertainty of population-scale transmission to be better modelled



Micro Model

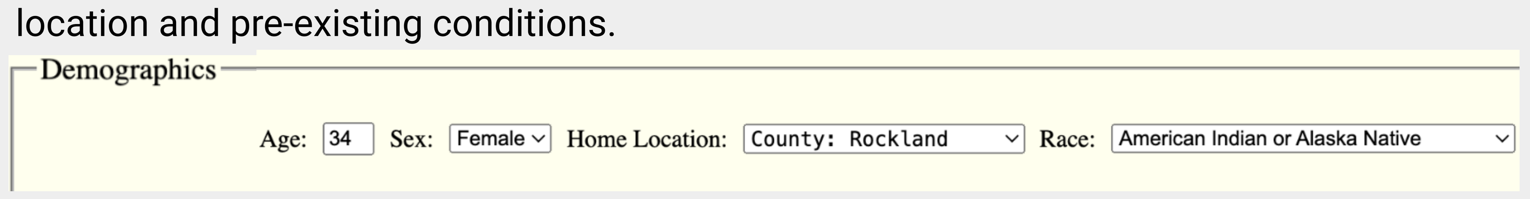
- The **Micro Model** focuses on assessing individual infection risk, adjusting the risk based on symptom testing and airborne transmission indoors. It uses personal data such as age, vaccination status, and chronic health conditions, along with a micro-mechanistic transmission model(s) (currently Bazant and Bush, 2021) to evaluate indoor airborne transmission. This data is then integrated with broader epidemiological models, such as the **macro-epidemiological model,** to provide a more accurate and personalized assessment of the risk of contracting COVID-19.



Generalized Risk Factor Framework

Generalized Risk Factor Framework is one component within **Pandemonium's** broader framework handling the storage, lookup, and processing of Risk and Protective Factors for users and models.

- Analyzed Factors: This module focuses on individual factors that influence the risk of infection, hospitalization, or death, including age, comorbidities, vaccination status, gender, ethnicity, location and pre-existing conditions.



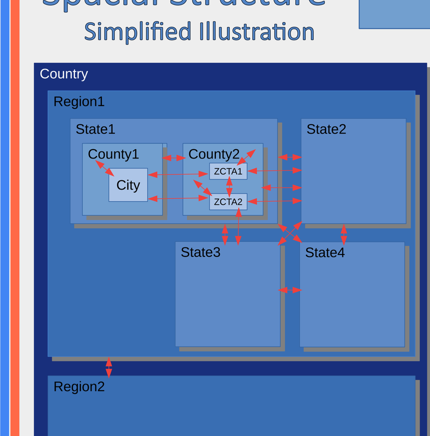
- The total relative risk compared to a standard baseline will be calculated with the user's given information.
- Converts between various forms, *e.g.*
  - Odds Ratio
  - Risk Ratio
  - Hazard Ratio
- Computes Relative Risk *vs.* Group/Local Population on-the-fly for Macro Model.

Future Plan for Pandemonium

1. Adapt **Pandemonium** to address future outbreaks of other vaccine-preventable diseases and emerging infections (like Influenza, RSV, and Monkeypox).
2. Improve **Vaccine Effectiveness model** by refining current models and incorporating data on individual vaccine responses, demographic information, and epidemiological trends to deliver more accurate risk assessments.
3. Develop and implement model of cases and/or active infections (**wastewater model**).
4. Scale the model globally through collaborations with researchers and public health organizations like the CDC and WHO, focusing on endemic infectious diseases (e.g. Malaria and Tuberculosis).
5. Incorporate **computational fluid dynamics** (CFD) simulations to further refine our understanding of airborne transmission in enclosed spaces.
6. Collaborate with the global research community and encourage wider adoption to improve emergency preparedness once the model **(1)** transitions to an **open-source** status starting Q4 2024.
7. Expand accessibility by releasing a mobile version and optimize the web interface.

Spacial Structure

Simplified Illustration

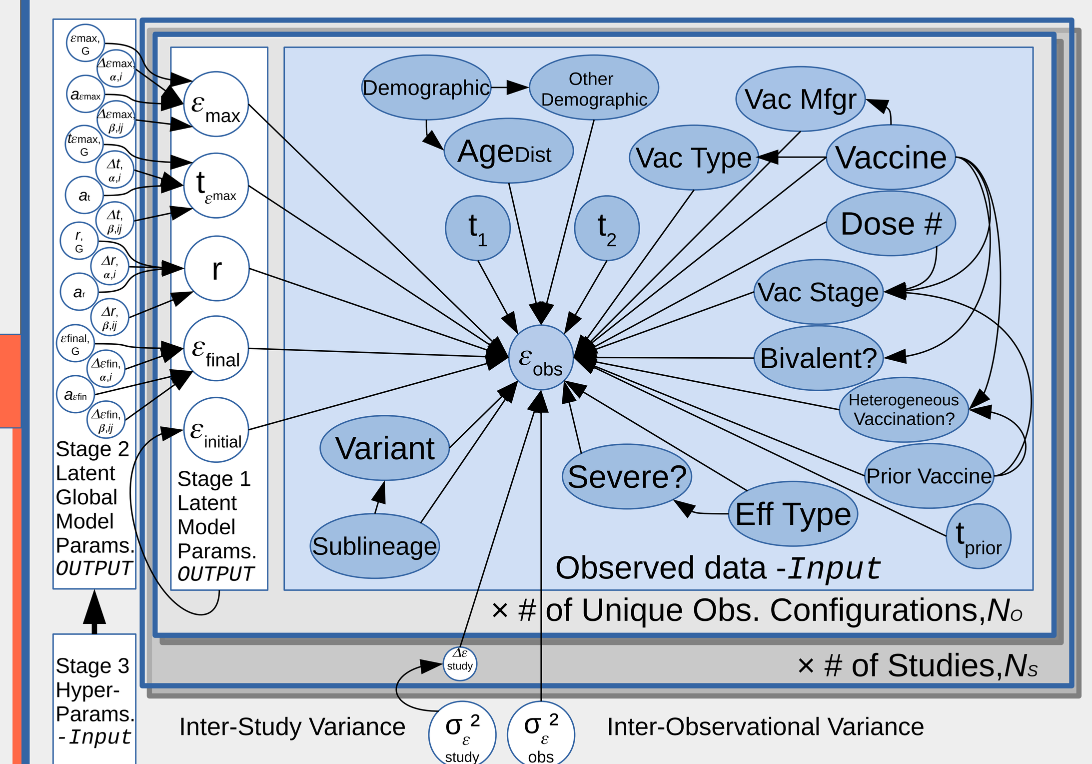


Blue Boxes : compartments
Blue arrows : flow through compartments
Teal Boxes : influencing factors resulting in parameter adjustments
Teal arrows : point to variable each alter
Purple circles : non-compartment latent parameters
Grey ovals : Observed data
Dashed gray arrows: Condition on observed data
Red arrows : Transmission, resulting in people moving from S to E
Yellow arrows : execution flow
Black arrows : data flow
p : response rate
μ : mortality rate
R : Effective Reproduction Number

Learn more about our innovative work

Scan the code for more info and to support our project!

Probabilistic Vaccine Effectiveness (VE) Model



ε: Vaccine Effectiveness
Eff Type: Effectiveness Type against different outcome
t_{prior}: Time between previous vaccination & latest vaccination
t_1, t_2: Time since last dose, lower & upper
r: exponential decay rate
Shaded circle: observed parameter
Clear circle: latent parameter

- Stage I: Together with sampled parameters, VE curve, average VE in a certain time interval and distributions of them are calculated.

Nonlinear Vaccine Effectiveness Model For each observation ε_{obs}, i.e. vaccine effectiveness data point,

$$\varepsilon_{\text{model}}(t) = f(t; \varepsilon_{\max}, t_{\varepsilon_{\max}}, r, \varepsilon_{\text{final}}, \varepsilon_{\text{initial}}) \quad (1)$$

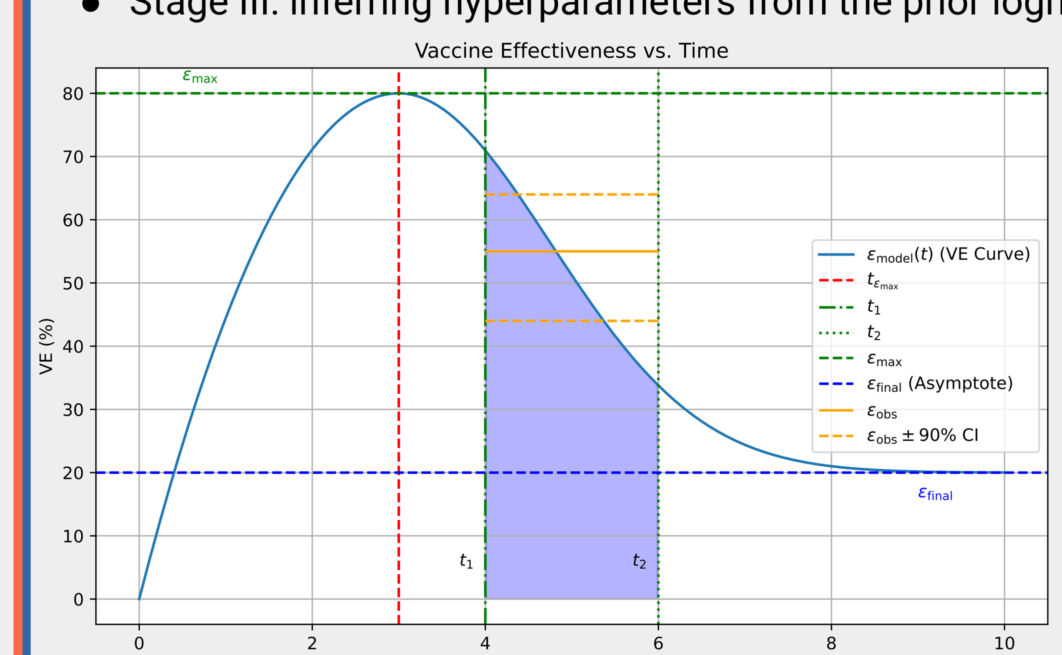
$$\varepsilon_{\text{avg}}(t_1, t_2) = \frac{\int_{t_1}^{t_2} \varepsilon_{\text{model}}\, dt}{t_2 - t_1} \quad (2)$$

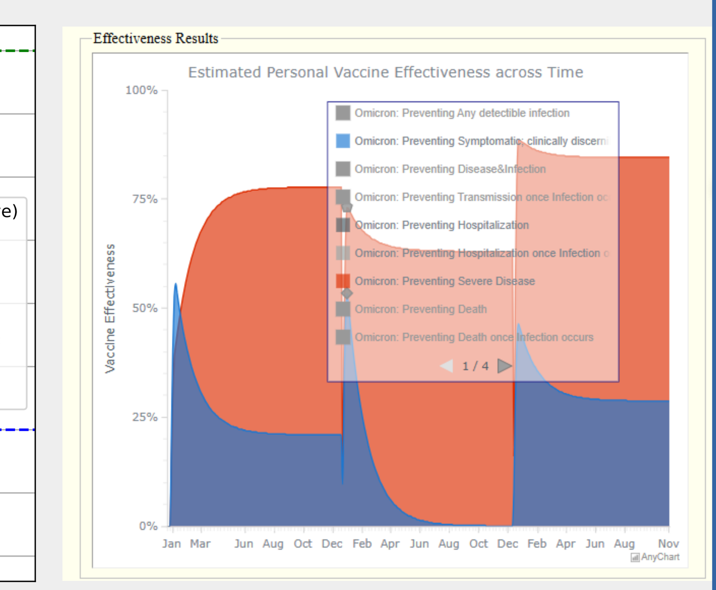
$$\varepsilon_{\text{avg}} \sim \mathcal{N}(\varepsilon_{\text{obs}} - \Delta\varepsilon_{\text{study}}, \sigma^2_{\varepsilon_{\text{obs}}}) \quad (3)$$

$$\varepsilon_{\text{obs}} \sim 1 - \text{LogNormal}(\log(1 - \mu_{\varepsilon_{\text{obs}}}) - \tfrac{1}{2}\sigma^2_{rr_{\text{obs}}}, \sigma^2_{rr_{\text{obs}}}) \quad (4)$$

$$\Delta\varepsilon_{\text{study}} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,\text{study}}) \quad (5)$$

- Stage II: inferring latent global model parameters based on the lognormal distribution with the hyperparameters from stage III, use them and demographic information (Age) to calculate the five parameters that could estimate the VE curve.
- Stage III: inferring hyperparameters from the prior lognormal/uniform distribution.





Our VE model infers VE curve model parameters that are most likely to fit the observed data. The VE model parameters can generate most-likely VE curves for observed & unobserved:

- Times
- Disease outcomes: severe, non-severe, hospitalization (observed only)
- Different vaccine combinations: vaccine manufacture, dosage number

Moreover, VE curve uncertainties are obtained from the parameter uncertainties.

Analysis:

We use both the method of MAP and MLE for estimation. MAP outperformed MLE by:

- Percentage of outliers (2.27% vs. 6.82%)
- Percentage of successfully estimated non-heterogenous vaccine (85.19% vs. 64.81%)
- Percentage of successfully estimated heterogenous vaccine (98.15% vs. 56.48%)