SISMID 2021: Mathematical Models of Infectious Disease

Day 2: Breakout Session on Heterogeneity and Herd Immunity

Population heterogeneity by age influences the herd immunity threshold

Science

REPORTS

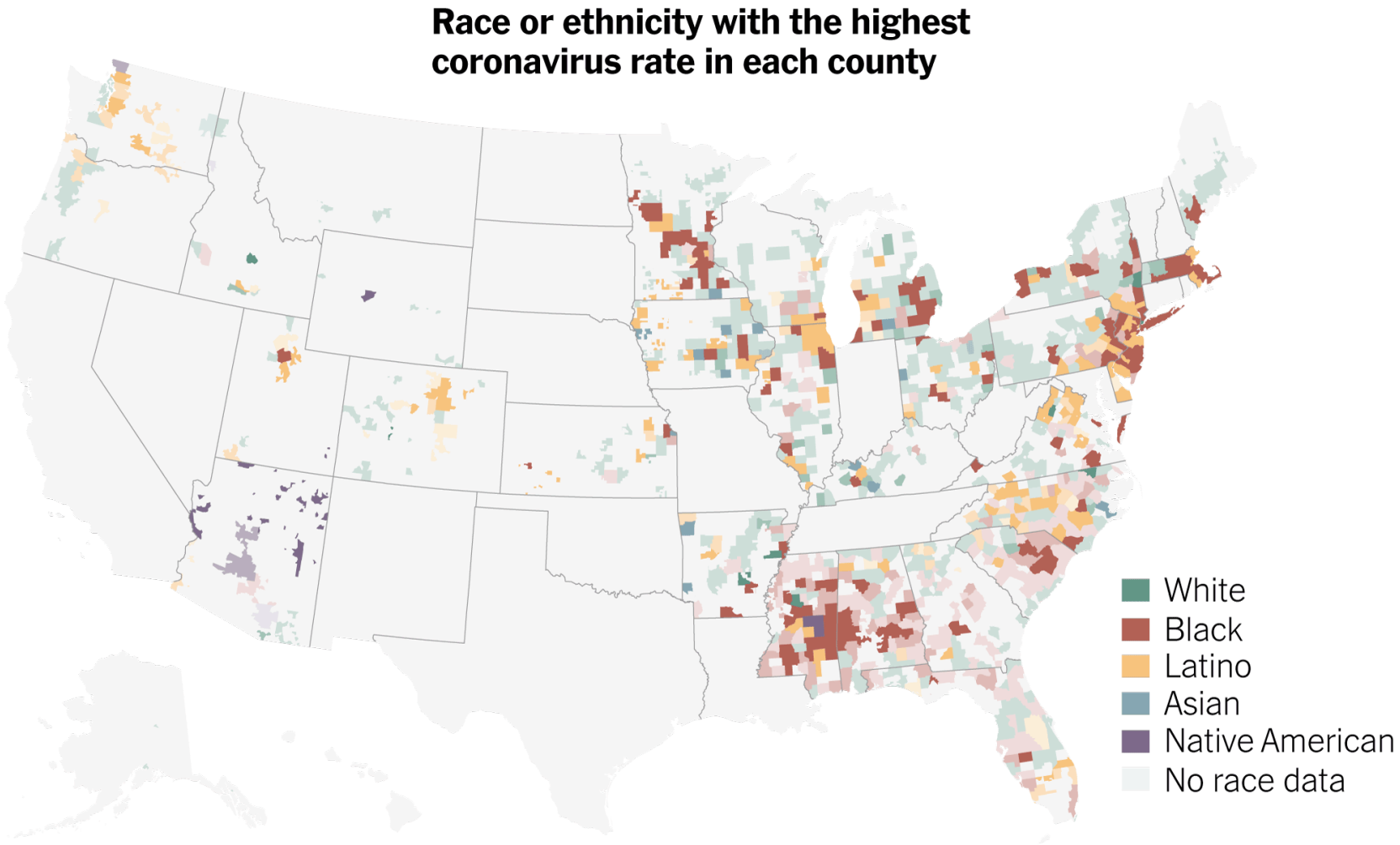
Cite as: T. Britton *et al.*, *Science* 10.1126/science.abc6810 (2020).

A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2

Tom Britton[1]*, Frank Ball[2], Pieter Trapman[1]

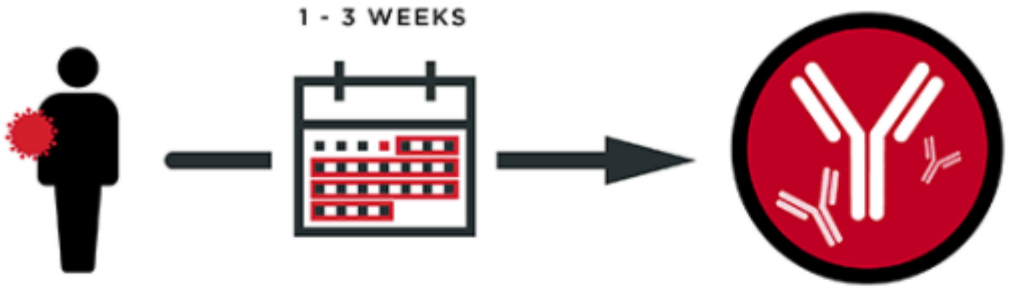
	$R_0 = 2.0$		$R_0 = 2.5$		$R_0 = 3.0$	
Population structure	h_D	h_C	h_D	h_C	h_D	h_C
Homogeneous	50.0	50.0	60.0	60.0	66.7	66.7
Age structure	46.0	50.0	55.8	60.0	62.5	66.7
Activity structure	37.7	50.0	46.3	60.0	52.5	66.7
Age and activity structure	34.6	50.0	43.0	60.0	49.1	66.7

Racial and ethnic disparities in COVID-19 incidence in the US



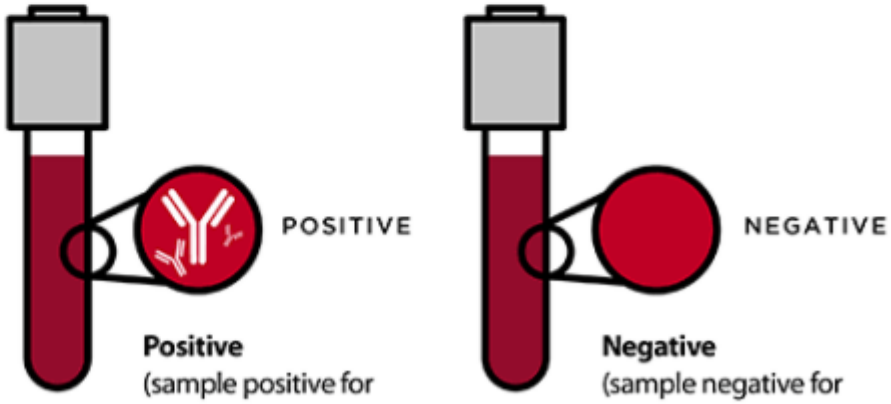
New York Times

Seroprevalence surveys allow us to estimate cumulative incidence of infection*



*Some people may take longer than 3 weeks to develop antibodies, and some people may not develop antibodies. It is currently unknown how long antibodies are detectable after infection.

A **positive** result from this test may mean that a person was **previously infected** with the virus.



Positive
(sample positive for SARS-CoV-2 antibodies)

Negative
(sample negative for SARS-CoV-2 antibodies)

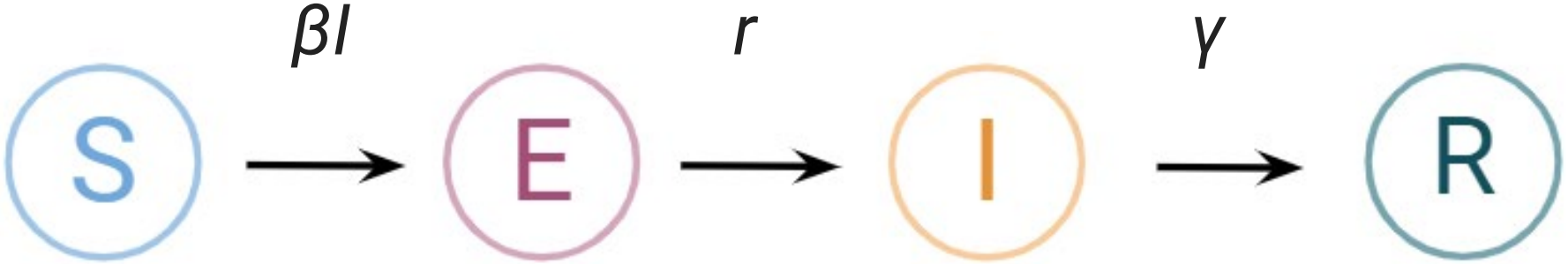
CDC

NYC serosurvey by Rosenberg et al.

- 100 grocery stores across New York state
- April 19-28, 2020
- 15,000 adults tested with information on race and ethnicity
- Overall cumulative incidence of 14%
- Limitation: single time point

Rosenberg et al., 2020

SEIR compartmental model for a homogeneous population



$$\frac{dS}{dt} = -\beta IS$$

$$\frac{dE}{dt} = \beta IS - rE$$

$$\frac{dI}{dt} = rE - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

susceptible

latently infected

infectious

recovered

(Seroprevalence data)

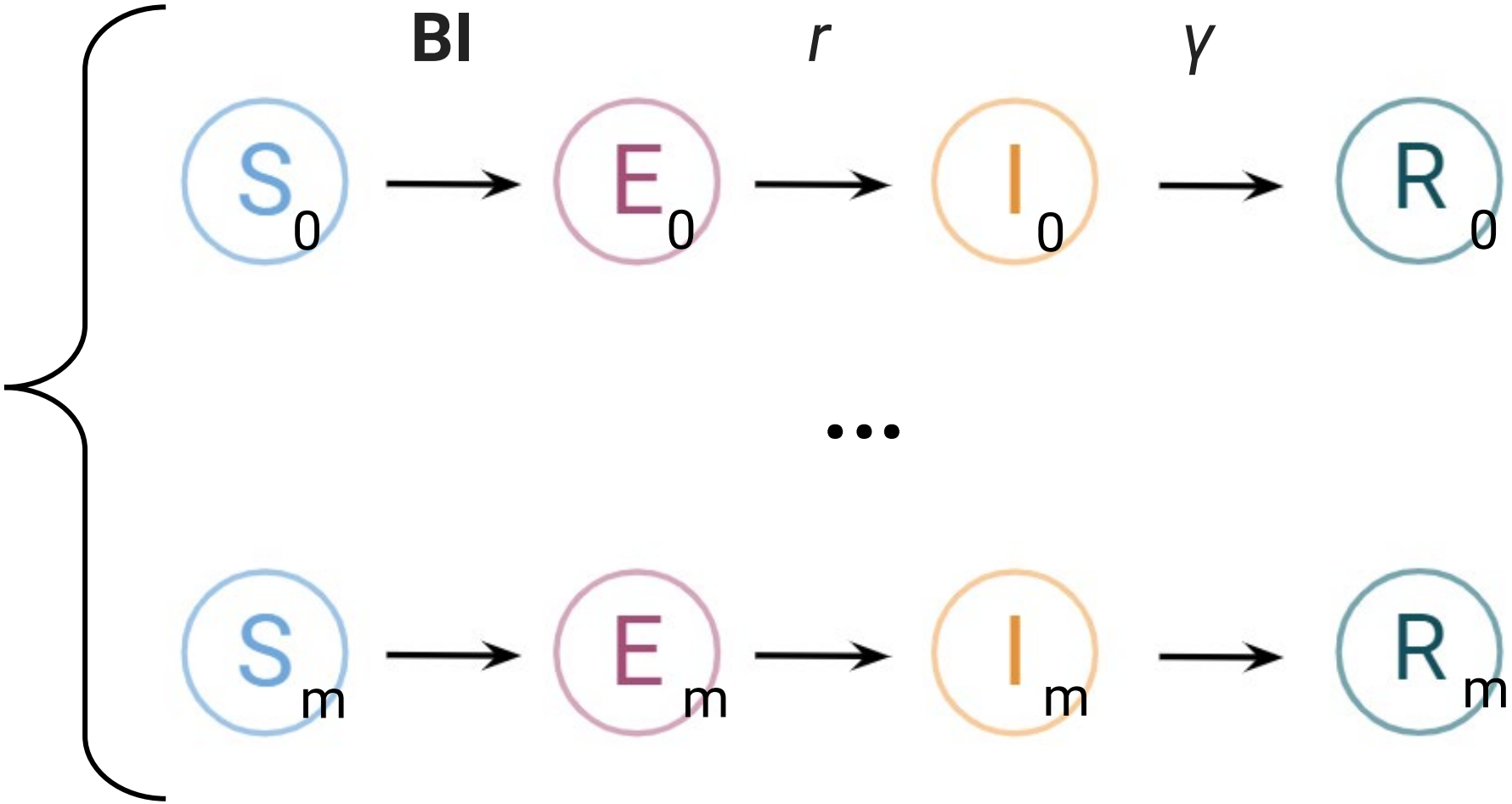
Structured models address heterogeneity by explicitly modeling the demographic subgroups

$$\frac{d\mathbf{S}}{dt} = -\mathbf{BI} \circ \mathbf{S}$$

$$\frac{d\mathbf{E}}{dt} = \mathbf{BI} \circ \mathbf{S} - r\mathbf{E}$$

$$\frac{d\mathbf{I}}{dt} = r\mathbf{E} - \gamma\mathbf{I}$$

$$\frac{d\mathbf{R}}{dt} = \gamma\mathbf{I}$$



Proportionate mixing assumes the contact rate between groups to be proportional to group activity levels (mean overall contacts per time)

$$\frac{d\mathbf{S}}{dt} = -\mathbf{BI} \circ \mathbf{S}$$

$$\frac{d\mathbf{E}}{dt} = \mathbf{BI} \circ \mathbf{S} - r\mathbf{E}$$

$$\frac{d\mathbf{I}}{dt} = r\mathbf{E} - \gamma\mathbf{I}$$

$$\frac{d\mathbf{R}}{dt} = \gamma\mathbf{I}$$

$$\mathbf{B} = q \begin{bmatrix} c_{0\leftarrow 0} & \cdots & c_{0\leftarrow 2} & \cdots & c_{0\leftarrow 4} \\ \vdots & & \vdots & & \vdots \\ c_{4\leftarrow 0} & \cdots & c_{4\leftarrow 2} & \cdots & c_{4\leftarrow 4} \end{bmatrix}$$

$$\beta_{i\leftarrow j} = q\frac{a_i a_j}{\sum_k a_k N_k}$$

Without contact survey data like POLYMOD, these assumptions are necessary

Assortative mixing partitions a fraction ε of contacts to be exclusively within group (with the rest proportionately distributed)

$$\beta_{i \leftarrow j} = (1 - \epsilon) * q \frac{a_i a_j}{\sum_k a_k N_k} + \epsilon * q \frac{a_i}{N_i}$$

More complexity generally increases model realism, but also parameter count

R_0, R_{eff}, and the herd immunity threshold (HIT) for heterogenous models

R_0 = dominant eigenvalue of $N\,\boldsymbol{B}\,/\,\gamma$

R_{eff} = dominant eigenvalue of $S(t)\,\boldsymbol{B}\,/\,\gamma$

HIT = fraction of population that is non-susceptible when $R_{eff} < 1$

Overview of breakout session project

- Fit transmission models structured by race and ethnicity to Long Island and New York City seroprevalence data
- Evaluate the impact of 1) different model structures and 2) variability in seroprevalence across groups on herd immunity thresholds and epidemic final sizes
- Think carefully through model assumptions, accuracy versus transparency versus resolution, etc.

Download code and overview

kevincma / sismid-2021-modeling-breakout

Unwatch

Code | Issues | Pull requests | Actions | Projects | Wiki | Security | Insights | Setti

main | 1 branch | 0 tags | Go to file | Add file | Code

kevincma Update README.md cc6a5eb 2 hours ago 12 commits

README.md	Update README.md	2 hours ago
SISMID_Day 2_Breakout_Slides.pdf	Create SISMID_Day 2_Breakout_Slides.pdf	3 days ago
SISMID_Day_2_Breakout_Overview....	update group numbers	2 hours ago
day-2-heterogeneiety-breakout.R...	Add comments	2 hours ago

Overview of the two main functions

- fit.model()
  - Input: **seroprevalence data**, demographics data, parameters (epsilon, latent period, infectious period)
  - Output: estimated activity level parameters
  - More on model fitting on Day 3

- run.structured.model()
  - Input: desired R0 value, demographics data, parameters (epsilon, latent period, infectious period, **activity levels**), initial conditions, timespan
  - Output: simulated epidemic trajectory

Schedule

- Work in breakout groups from now until 2:10 PM Pacific / 5:10 PM Eastern (~1 hour 10 minutes)
  - Read through PDF and code (~20 minutes)
  - Work on questions as a group (~45 minutes)
  - Finalize solutions and figures (if relevant) for your assigned question + be ready to present (~last 5 minutes)

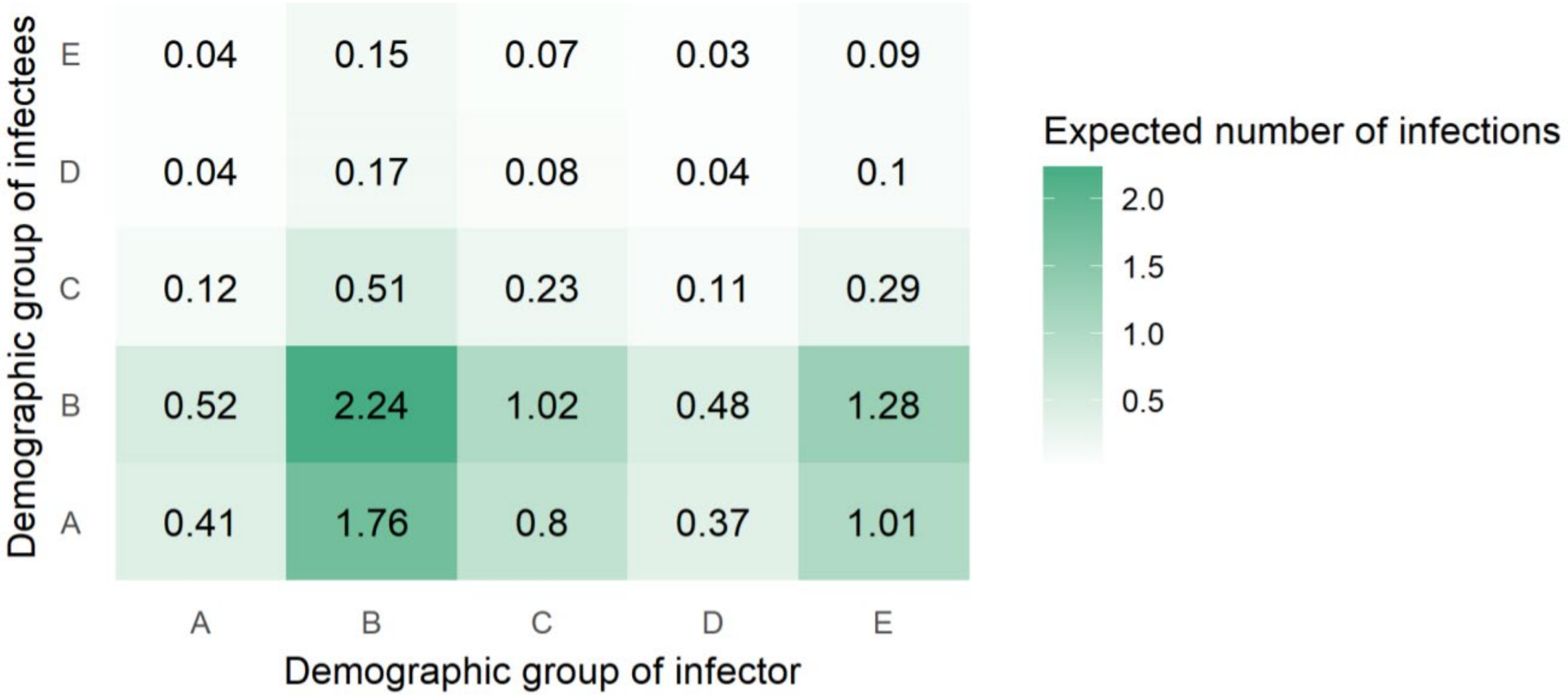
- Discuss in big group for the last 20 minutes of class

Solutions to breakout session

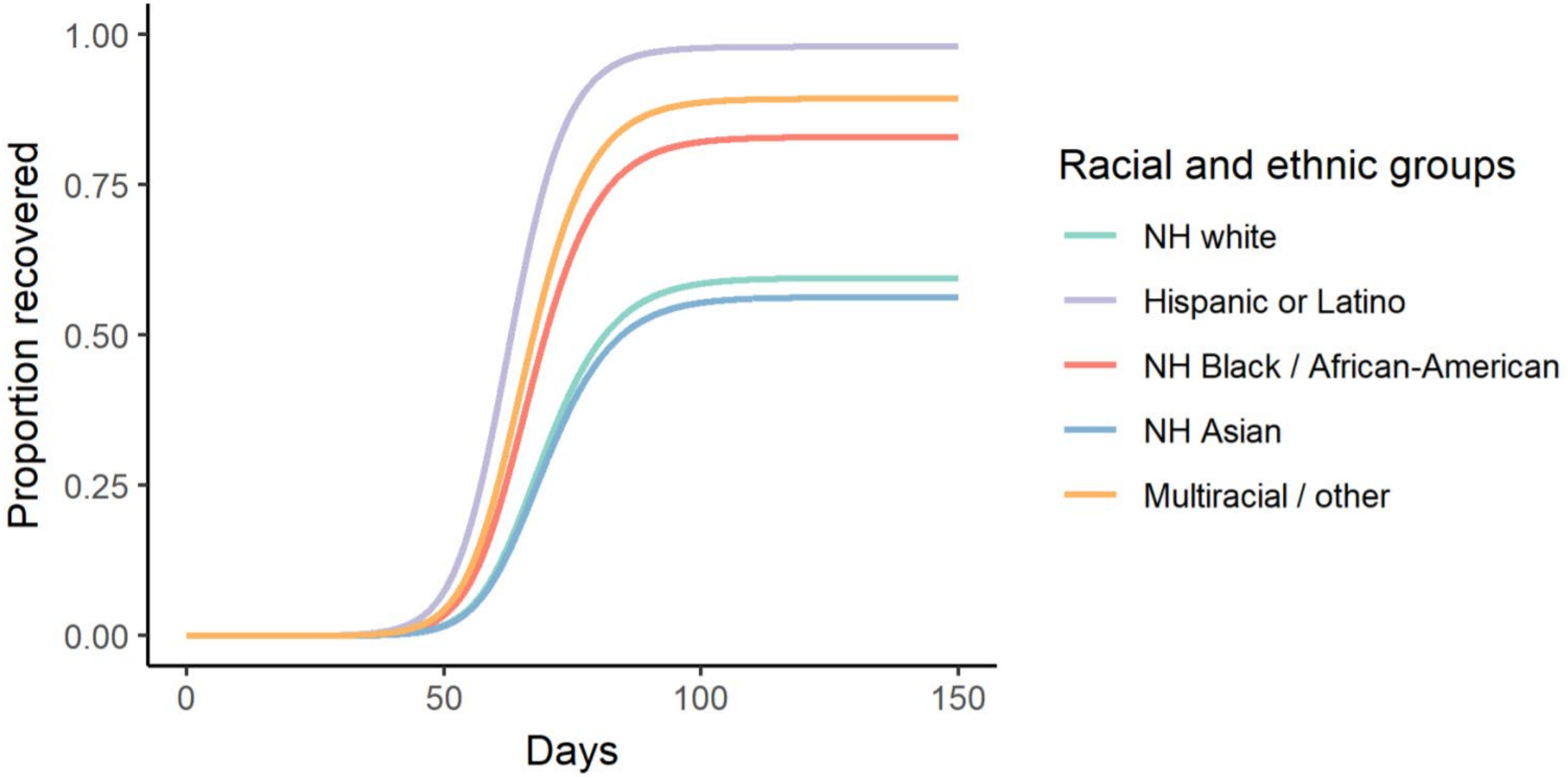
A) Long Island proportionate mixing model outputs

```
[1] "Normalized fitted values: 1 4.31 1.956 0.916 2.475"
$`Final epidemic size`
[1] 0.6932
$HIT
[1] 0.3977
$`Cumulative incidence by group at the HIT`
[1] 0.2861 0.7660 0.4828 0.2657 0.5658
```

A) Next-generation matrix



A) Epidemic trajectory plot (cumulative incidence by group)



A) Model assumptions

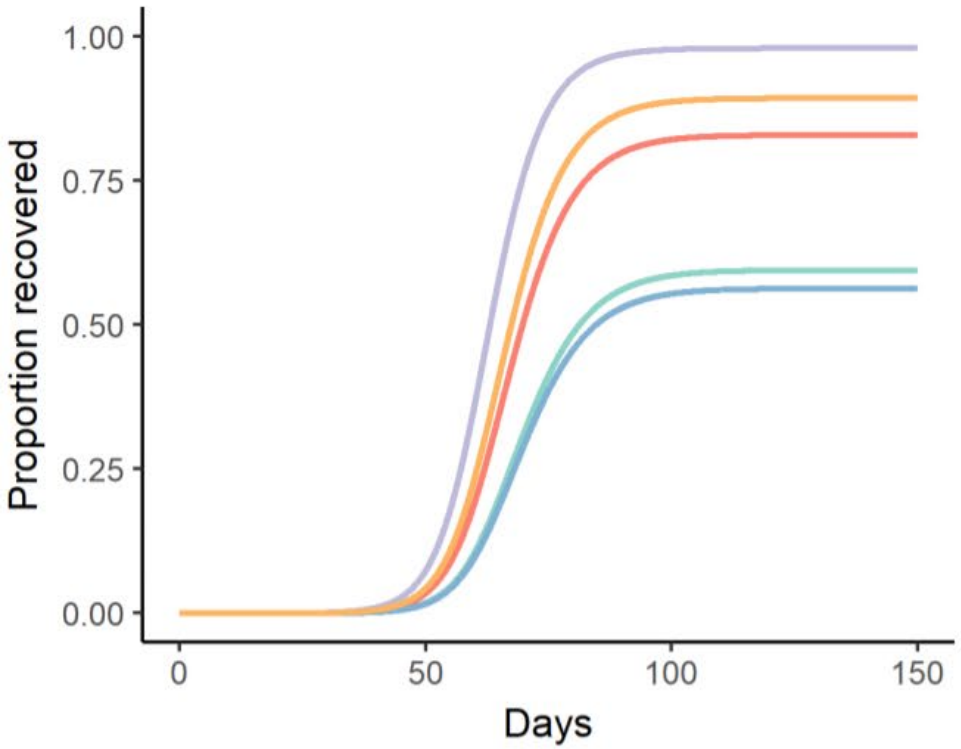
- Transparent and flexible, but with reduced accuracy and resolution compared to more complex models
- Seropositivity implies complete immunity (full protection from re-infection) and immunity does not wane; seroprevalence survey is unbiased
- Variability in exposure is modeled as a function of only one demographic characteristic
- Etc.

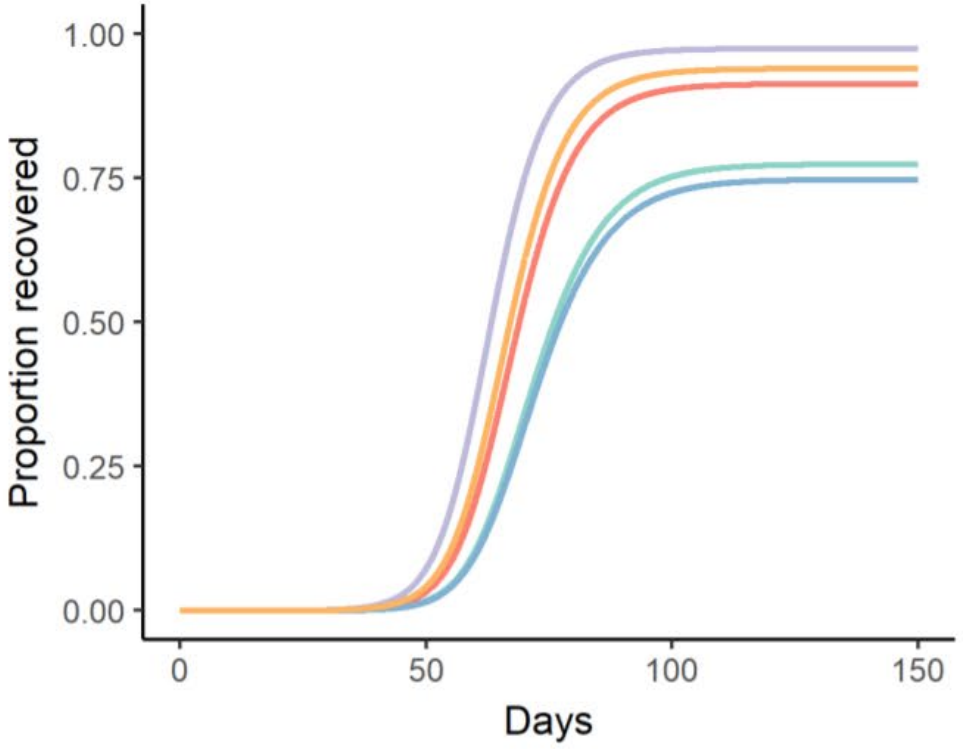
B) Assortative mixing model outputs

```
[1] "Normalized fitted values: 1 2.196 1.516 0.941 1.714"
$`Final epidemic size`
[1] 0.8268
$HIT
[1] 0.4955
$`Cumulative incidence by group at the HIT`
[1] 0.3875 0.8342 0.6099 0.3615 0.6871
```

B) Comparing next-generation matrices (proportionate – left vs assortative mixing – right)

B) Comparing epidemic trajectories (proportionate – left vs assortative mixing – right)





Proportionate mixing

Assortative mixing (ε=0.5)

C) NYC model outputs (proportionate mixing – left, assortative mixing – right)

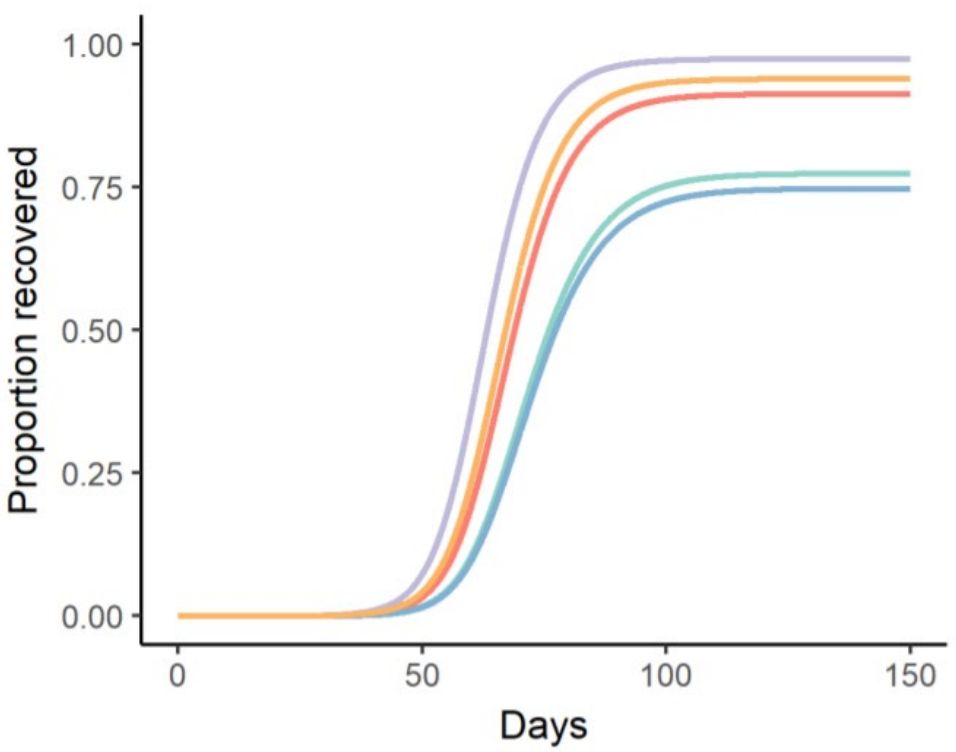
```
$`Final epidemic size`
[1] 0.8727

$HIT
[1] 0.5759
```

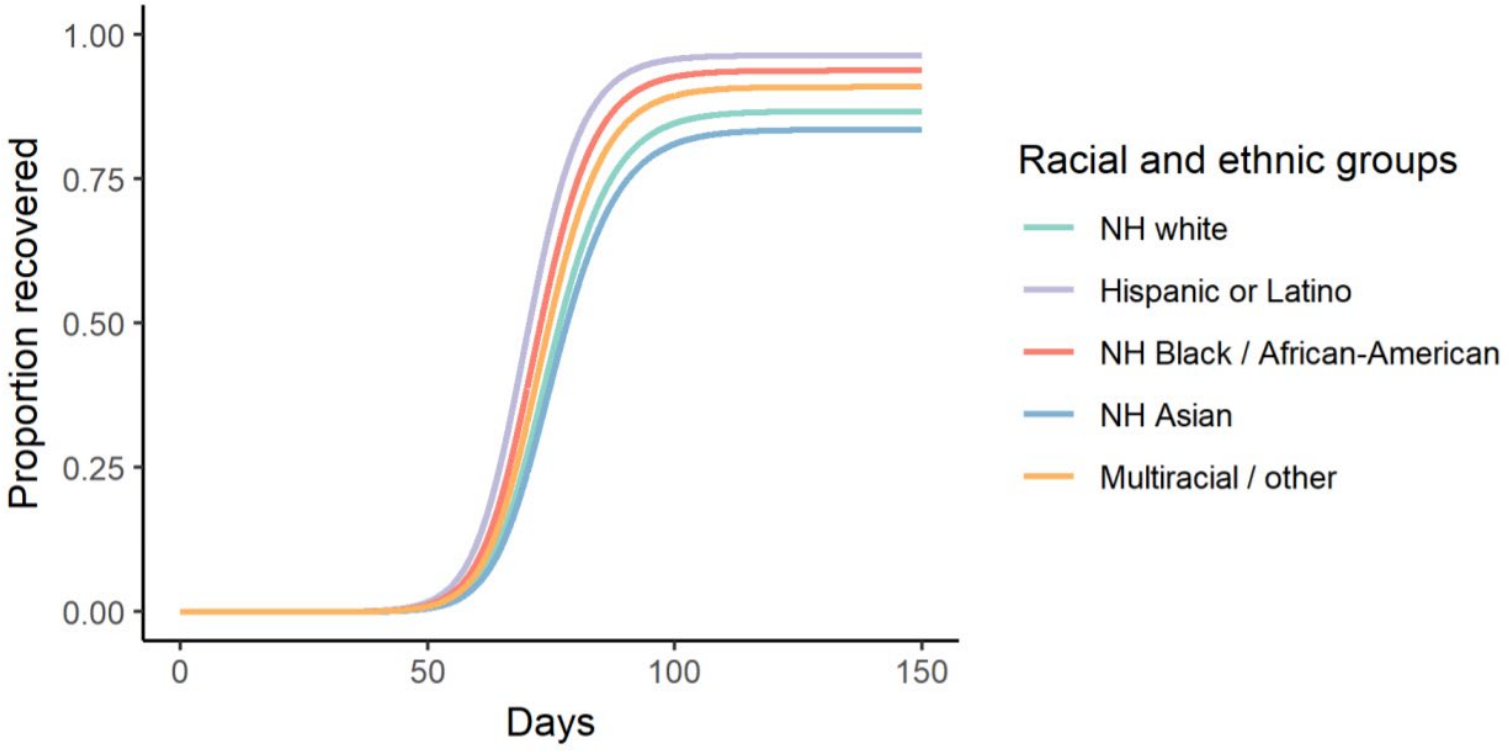
```
$`Final epidemic size`
[1] 0.9085

$HIT
[1] 0.6105
```

C) Epidemic trajectories for assortative mixing models in Long Island (left) and NYC (right)



Long Island



NYC

C) Long Island versus NYC seroprevalence data

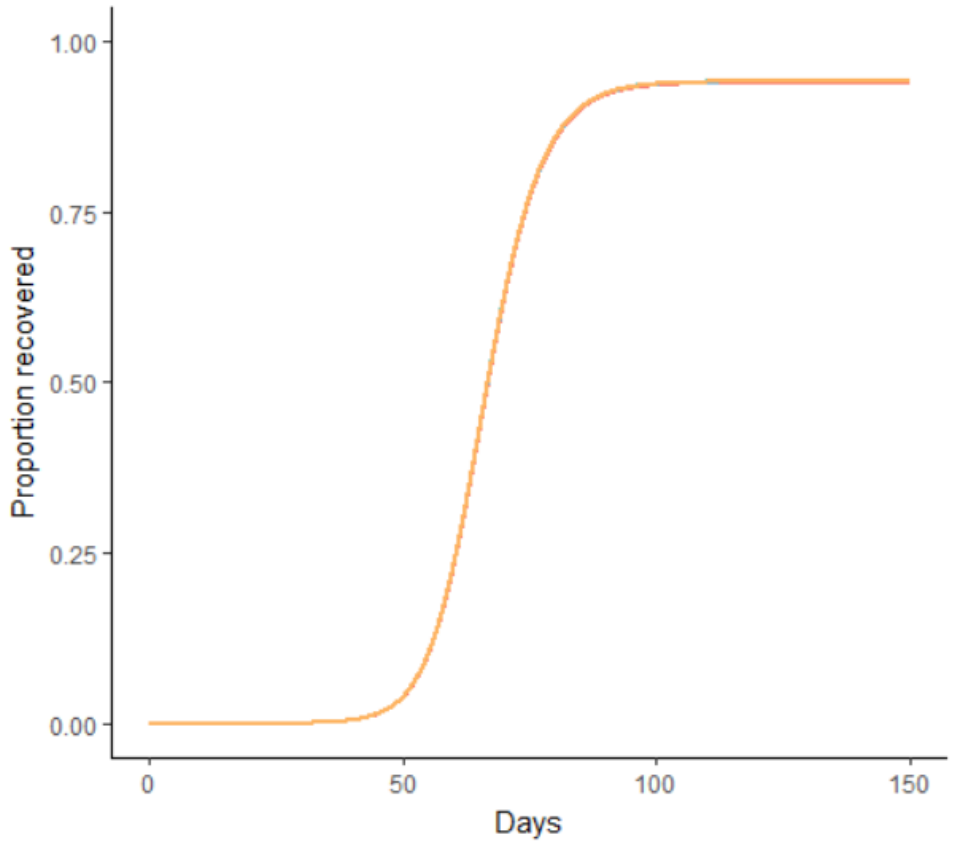
	Long Island	NYC
Non-Hispanic whites	0.087	0.166
Hispanics or Latinos	0.320	0.330
Non-Hispanic Blacks	0.158	0.252
Non-Hispanic Asians	0.084	0.145
Multiracial / other	0.207	0.204
Coefficient of variation (standard deviation / mean)	**0.57**	**0.33**

D) Model outputs for a homogeneous population

```
activities.homogeneous <- fit.model(pop=pop.li, pop.p=pop.p.li, sero.p=rep(0.2,5),
    sero.tested=sero.tested.li,
                                epsilon=0, r.parm=r.parm, g.parm=g.parm)
results.homogeneous <- run.structured.model(R0.value=default.R0, pop=pop.li,
    pop.p=pop.p.li, epsilon=0,
    g.parm=g.parm, fitted.vars=activities.homogeneous)
print(results.homogeneous)


### Output
[1] "Normalized fitted values: 1 0.996 0.989 0.999 1"
$`Final epidemic size`
[1] 0.9414


$HIT
[1] 0.6666
```

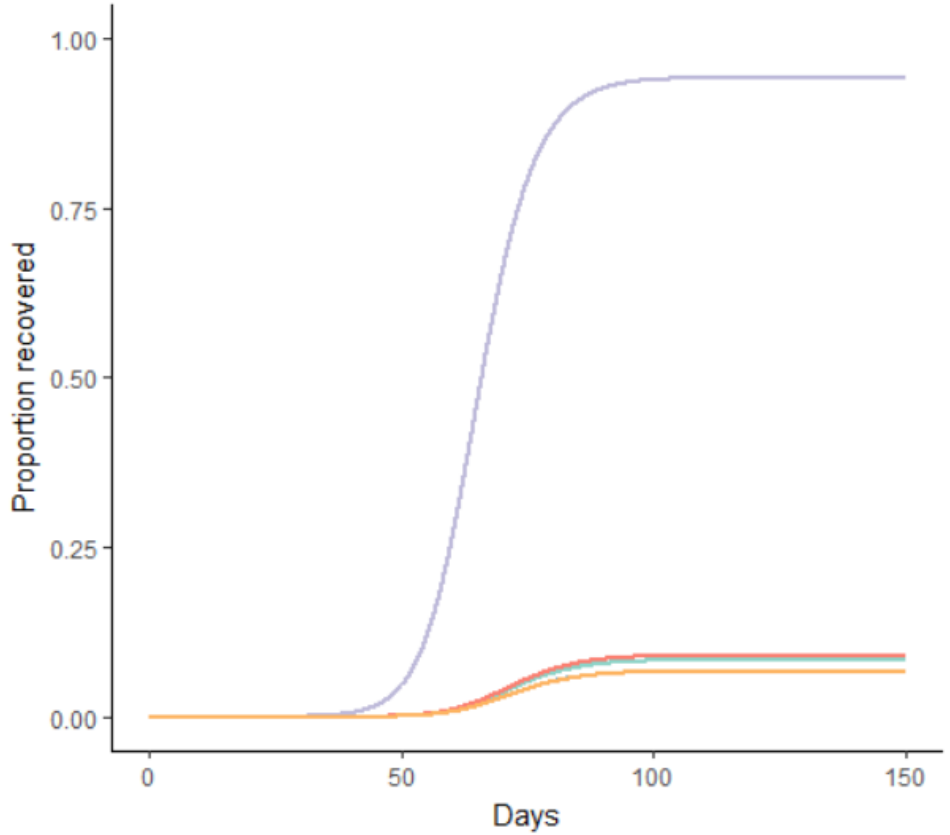


D) Model outputs for a highly heterogeneous population

```
activities.heterogeneous <- fit.model(pop=pop.li, pop.p=pop.p.li,
    sero.p=c(0.05,0.8,0.05,0.05,0.05), sero.tested=sero.tested.li,
                                epsilon=0, r.parm=r.parm, g.parm=g.parm)
results.heterogeneous <- run.structured.model(R0.value=default.R0, pop=pop.li,
    pop.p=pop.p.li, epsilon=0,
    g.parm=g.parm, fitted.vars=activities.heterogeneous)
print(results.heterogeneous)

### Output
$`Final epidemic size`
[1] 0.2431

$HIT
[1] 0.1506
```



Summary of questions A through D

- Both variability in the data and the choice of model affects HIT and final epidemic size estimates
- All else equal (holding R0 and model type constant), increased variability in the seroprevalence data will generally reduce HITs and final epidemic sizes
- All else equal (holding R0 and data / location constant), assuming proportionate mixing will reduce HITs and final sizes versus assortative mixing models

E) Including non-pharmaceutical interventions (NPIs)

- Scale the transmission rate by a factor beginning when 5% cumulative incidence in the population was reached – representing an established and expanding epidemic – for a variable duration
- Extensions: allow linear increases / decreases in NPIs, parameterize using mobility data, etc.