

Introduction to Data Science for Public Health

Michelle Campbell, MSHI

Director, Center for Data Modernization and Informatics,
Washington State Department of Health

Ian Painter, PhD

Senior Epidemiologist
Washington State Department of Health



Using Zoom Q&A



If you have a question during the presentation, please click the **Q&A icon** in the Zoom toolbar to open your Q&A Pod.

About the Public Health Data Learning Center



Public Health
DATA LEARNING CENTER



SCHOOL OF PUBLIC HEALTH
UNIVERSITY *of* WASHINGTON



Introduction to Data Science

Data Modernization Initiative Updates



Using Pandemic Data Modeling to Inform Public Health Decisions

Ian Painter Ph.D | Senior Epidemiologist

Washington State Department of Health



Outline

- 1. A brief history**
- 2. Primary model and data sources**
 - i. ID modeling workflow
 - ii. SEIR model
- 3. How we use modeling**
 - i. Modeling to inform policy decisions
 - ii. Implementing health equity into modeling



Question for the Audience

During the COVID-19 pandemic, did you work with any of the following types of partners on data science projects (any project involving collecting, processing, analyzing or interpreting data)?

- A. Academic groups
- B. Other public health agencies
- C. Private industry
- D. Not-for-profit organizations
- E. Did not partner
- F. Other (Type in chat)



A brief history

MARCH 15, 2020

The Seattle Times
Sunday
WINNER OF 10 PULITZER PRIZES

MOSTLY SUNNY,
HIGH CLOUDS
High, 46. Low, 35. > 88
seattletimes.com/weather

\$3.00
SEATTLETIMES.COM

'I KIND OF WANT TO CRY' | Life goes on despite the coronavirus, but it's muted, as schools, courts, public offices, landmarks, cultural institutions and many businesses shut their doors.

SILENCE IN SEATTLE



Bellevue scientists modeling the size of outbreak

By SANDI DOUGHTON
Seattle Times staff reporter

When Gov. Jay Inslee announced a ban on large gatherings last week and warned schools to prepare for extended closures, he was standing next to a chart that showed the consequences of inaction: a relentless, upward trajectory in new coronavirus infections with 25,000 cases by early April and 400 people likely to die in King and Snohomish counties alone.

Those sobering projections were generated by a little-known team of local researchers who until recently focused mainly on tuberculosis, malaria and other scourges of the developing world. But as they watched the novel coronavirus spread around the globe and land in their own backyard, the epidemiologists and computer whizzes at the Institute for

Timeline: Pre-Pandemic - Present



Before COVID-19 pandemic:
Early IDM work



April 2020:
**COVID-19 Modeling
Working Group**



Present:
Model Updates



March 2020:
**IDM / DOH
Governor Briefings**



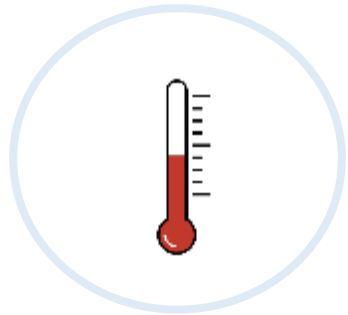
2021:
Hand off to DOH

The image features a person's hands typing on a laptop keyboard. The scene is overlaid with various digital data visualizations. On the left, there is a vertical list of binary code (0s and 1s) in green and blue. In the center, a network diagram shows a central node labeled '100111' with lines radiating outwards. On the right, there is a bar chart with blue and yellow bars, and a line graph with a blue and yellow wavy line. In the foreground, a circular graphic contains the text '80%'. To the right of the keyboard, there are three horizontal bars with labels '88%', '71%', and '46%'. The overall color scheme is dark blue and black, with highlights from the data visualizations.

Primary model and data sources

ID modeling workflow

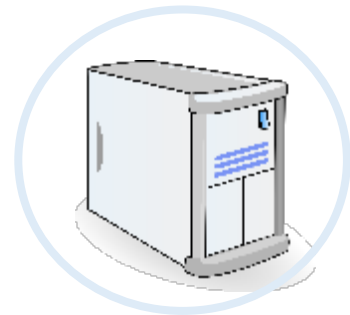
Data collection



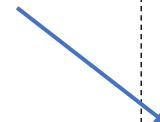
Sick people



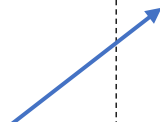
Health care system



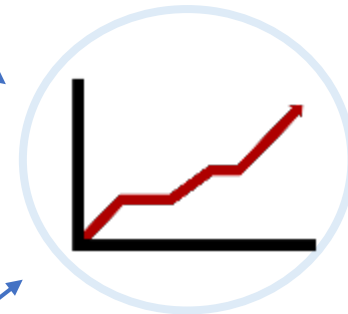
Data aggregation



Other data sources



Modeling

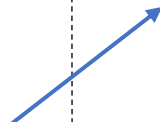


ID model

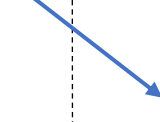
Communication



Decision makers



General public



What do we *actually* measure?

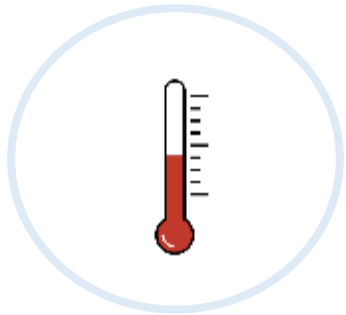
What we want to know	What we can measure
How many people are infected?	
How much transmission is happening?	
How much burden is COVID-19 putting on people?	
How much burden is COVID-19 putting on our health care system?	
What might happen going forward?	
What can we do about this?	

What do we *actually* measure?

What we want to know	What we can measure
How many people are infected?	<ul style="list-style-type: none">How many cases are occurring
How much transmission is happening?	<ul style="list-style-type: none">Very difficult to directly measure
How much burden is COVID-19 putting on people?	<ul style="list-style-type: none">Good measurement on severe health outcomes (hospital admissions and deaths from COVID-19)Less information on economic, educational, and psychological burdenLess information on who the burden is falling on
How much burden is COVID-19 putting on our health care system?	✓
What might happen going forward?	✗
What can we do about this?	✗

ID modeling workflow

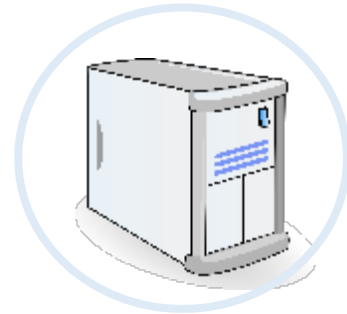
Data collection



Sick people



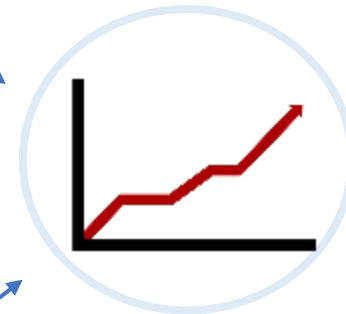
Health care system



Data aggregation



Modeling



ID model

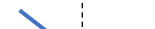
Communication



Decision makers



General public



What we measure:

- Cases
- Hospital admissions
- Deaths

What we want to know:

- Infections
- Transmissions
- What might happen



Other data sources



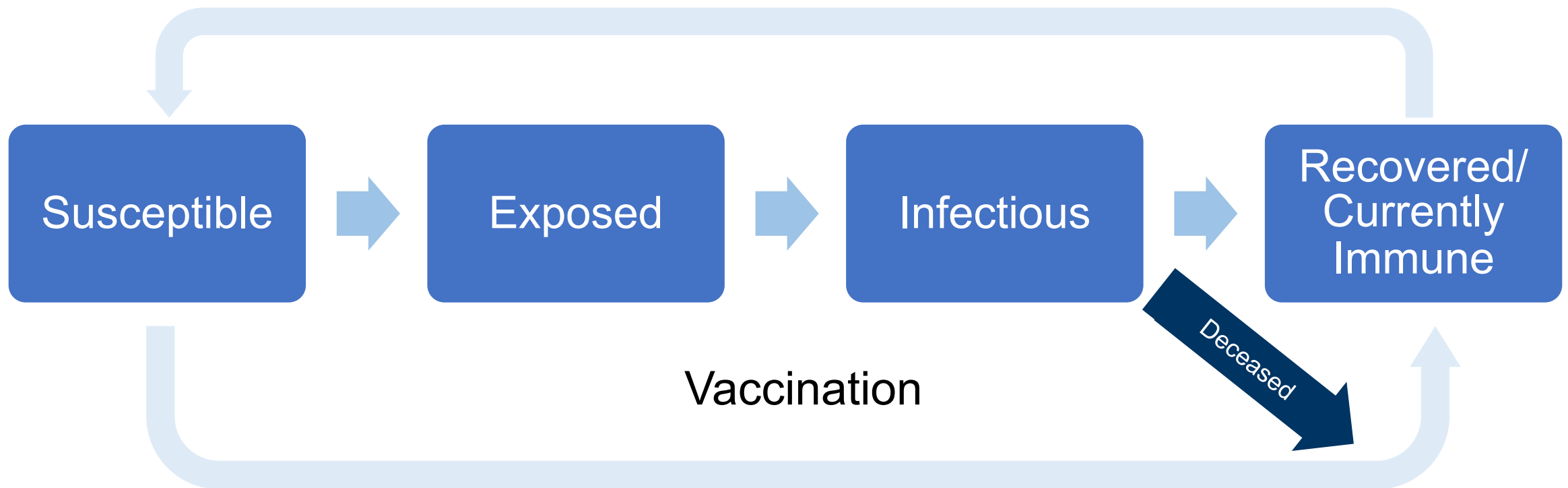
Washington state data sources

- Tests and cases (WDRS)
- Hospitalizations (WDRS, RHINO, WA Health)
- Deaths
- Vaccinations – ~95% complete
- Seroprevalence (CDC)
- Sequence data
- Wastewater data

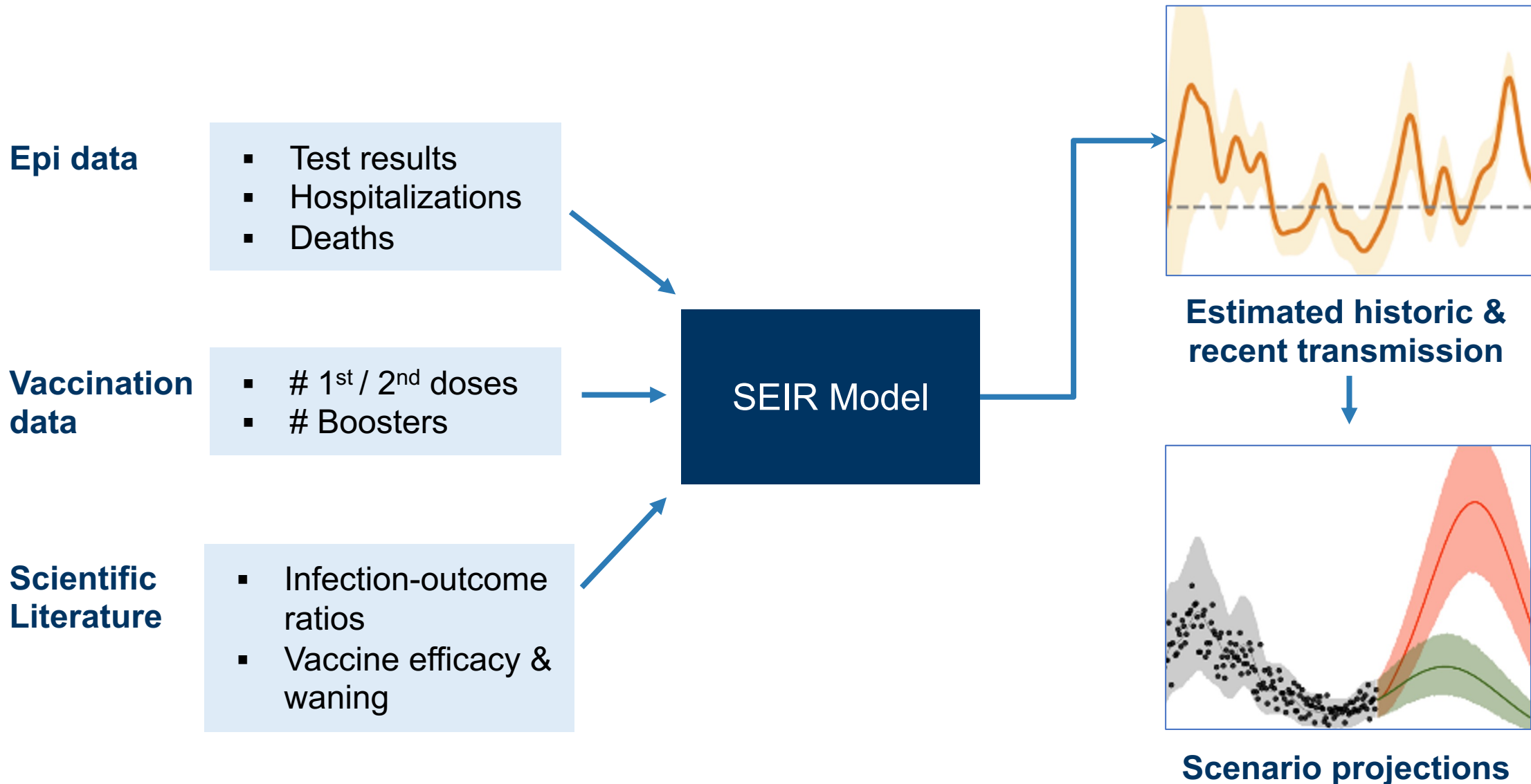


SEIR model

Immune escape / waning

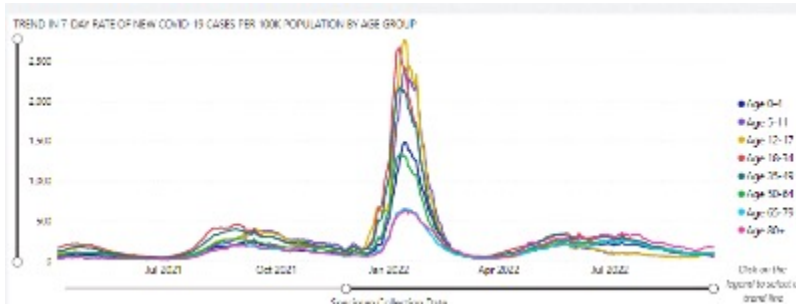


SEIR modeling workflow

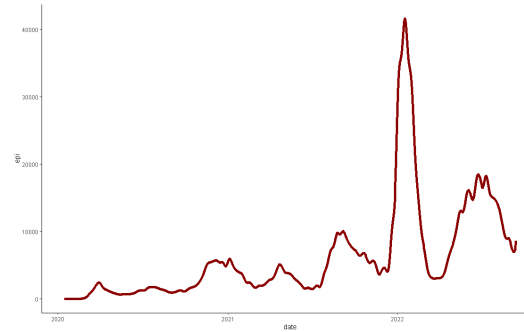


Analytic process

Hospital admissions

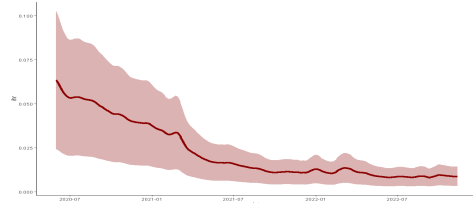


Infection epi-curve

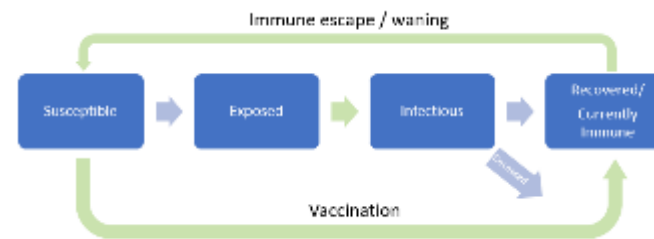


Case count curves
Serology data

Infection hospitalization rate (IHR)



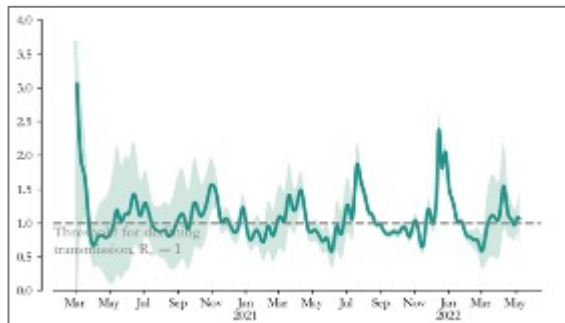
SEIR model



Model parameters

- Incubation period
- Infectious period
- Waning of immunity(s)
- Immunity(s) from vaccination

Transmission
Prevalence
Immunity



Infection hospitalization rate

Baseline IHR:

Variant effects



Immunity effects

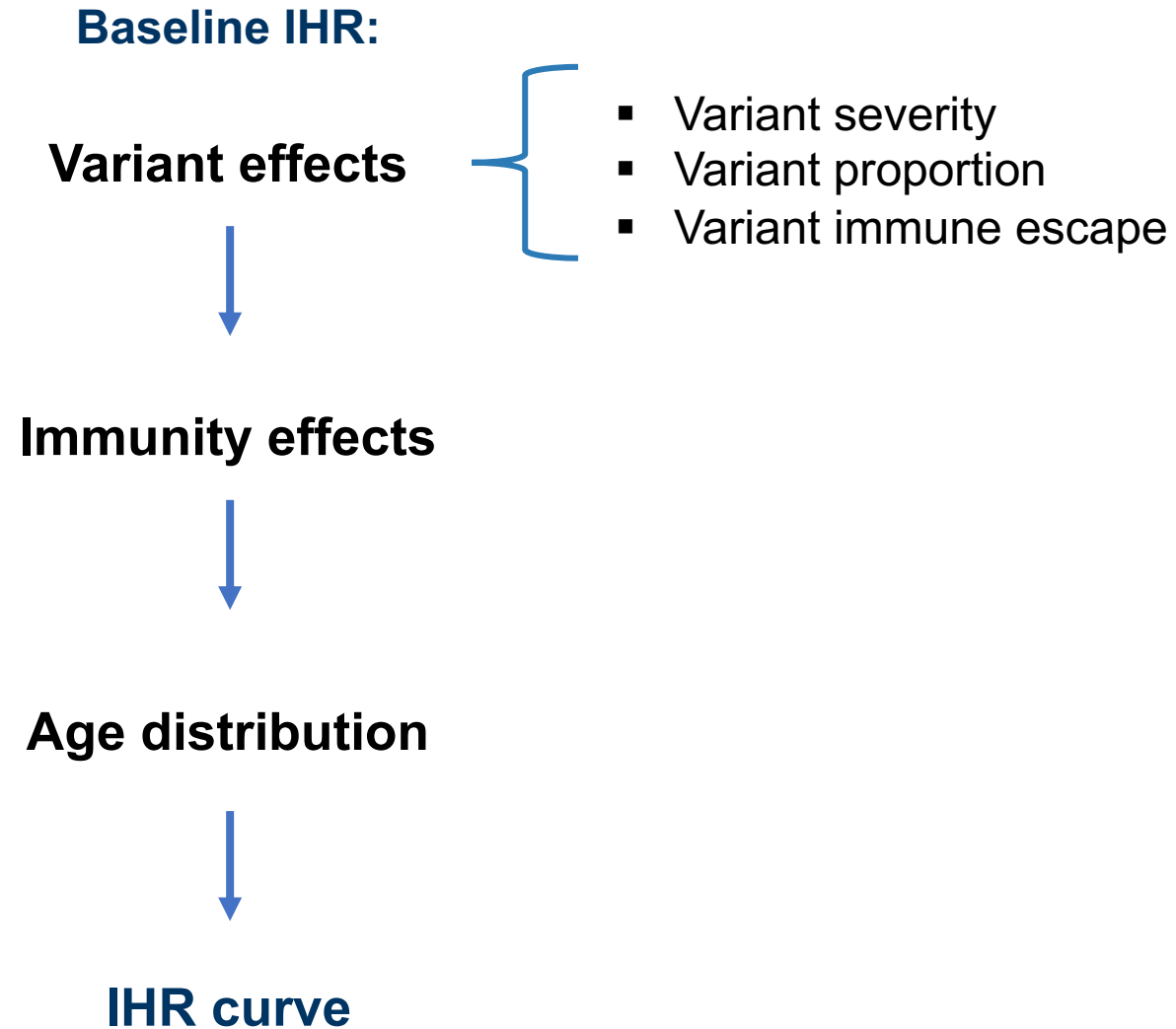


Age distribution

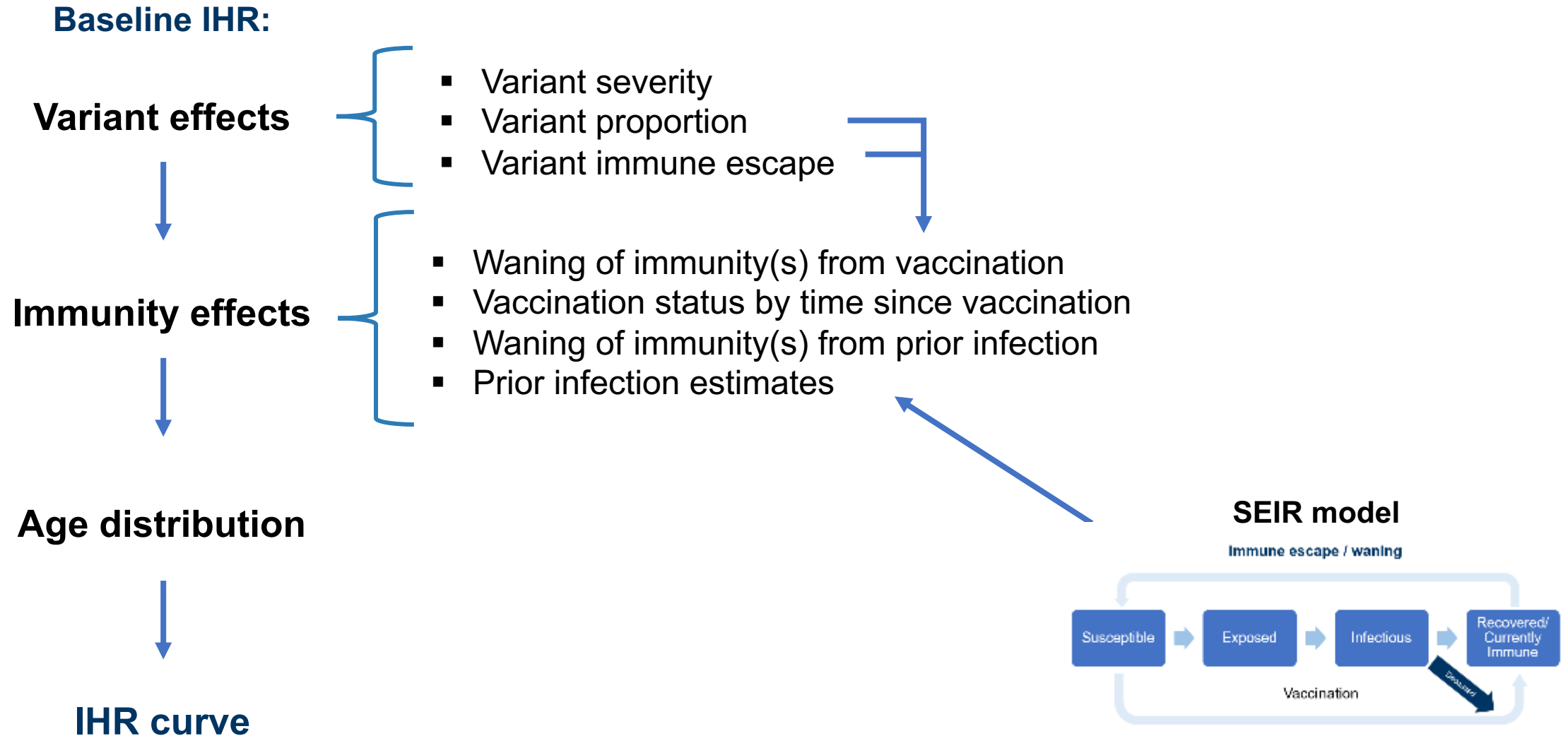


IHR curve

Infection hospitalization rate



Infection hospitalization rate



Infection hospitalization rate

Baseline IHR:

Variant effects

- Variant severity
- Variant proportion
- Variant immune escape

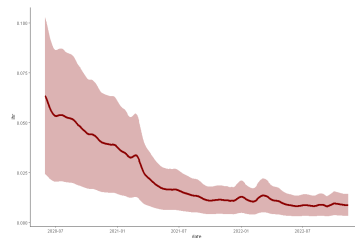
Immunity effects

- Waning of immunity(s) from vaccination
- Vaccination status by time since vaccination
- Waning of immunity(s) from prior infection
- Prior infection estimates

Age distribution

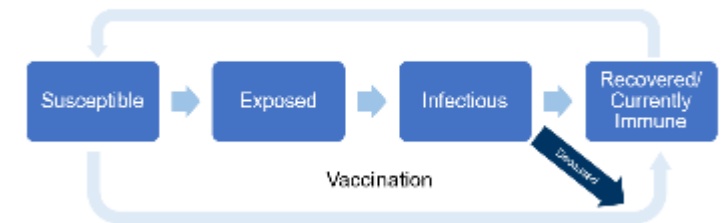
- Population estimates
- Case counts
- Serology

IHR curve



SEIR model

Immune escape / waning





Question for the Audience

What barriers do you face in working with outside partners on data science projects during public health emergencies?

- A. Access to data
- B. Access to suitable partners
- C. Data sharing challenges
- D. Communicating needs and capabilities
- E. Other (Type in chat)

A person wearing a light blue button-down shirt is seated at a desk, working on a laptop. The laptop screen displays a glowing blue flowchart with several rectangular nodes and connecting arrows. The flowchart is semi-transparent and appears to be floating above the laptop. The background is a soft, out-of-focus blue. A dark horizontal bar is overlaid across the middle of the image, containing the text 'How we use modeling' in white.

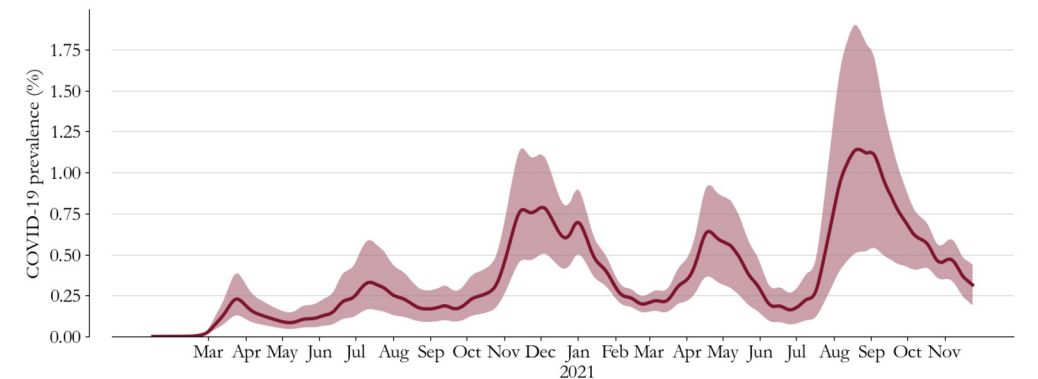
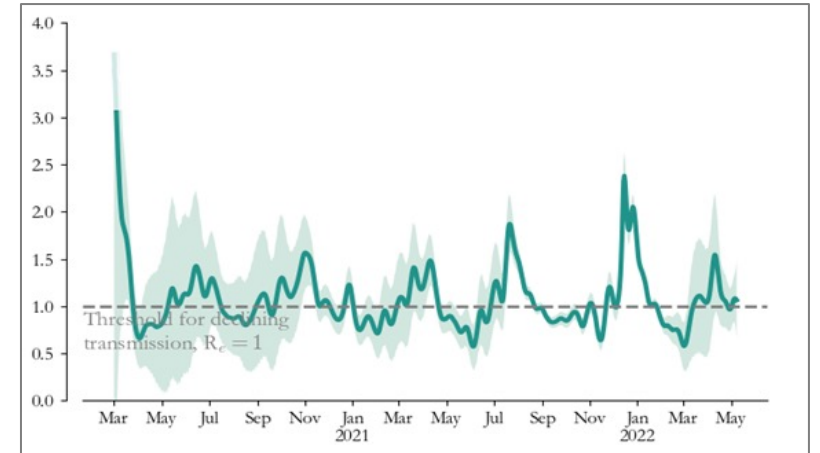
How we use modeling

Case 1: Current situation

Population diagnostic machine:

What is currently happening with disease transmission?

- R effective
- Prevalence
- Immunity



Case 2: Scenario projections

- Projections require additional assumptions:
 - Future changes in transmission (Seasonal, Behavioral)
 - Future vaccination rates
 - New variant properties
- Projections based on assumptions about the effect of specific scenarios on future transmission
- Transmission assumptions feed into the model to project future admissions



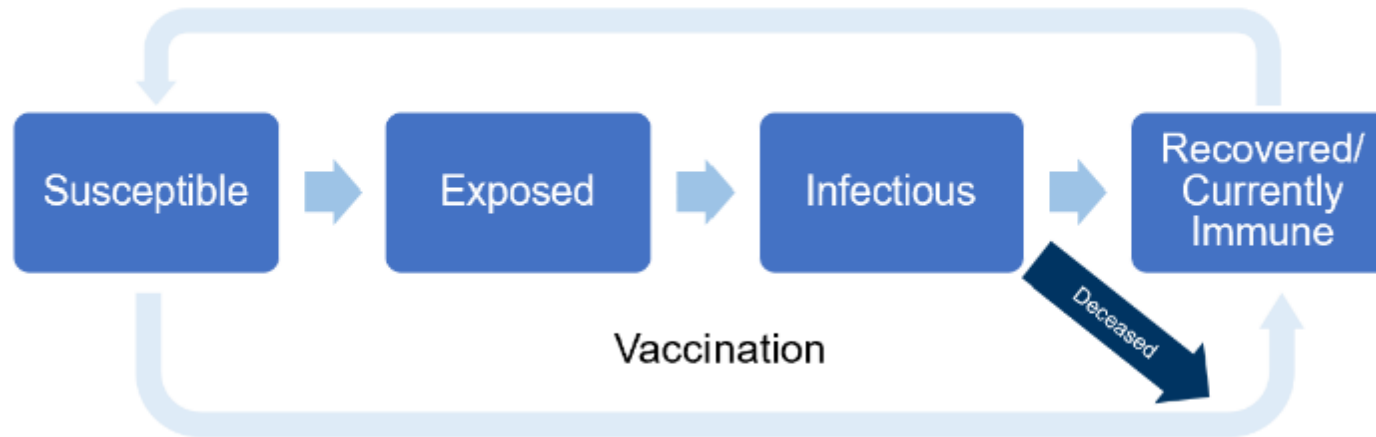
Case 2: Scenario projections

Estimation

Hospital Admissions → Infection Epi-curve

IHR

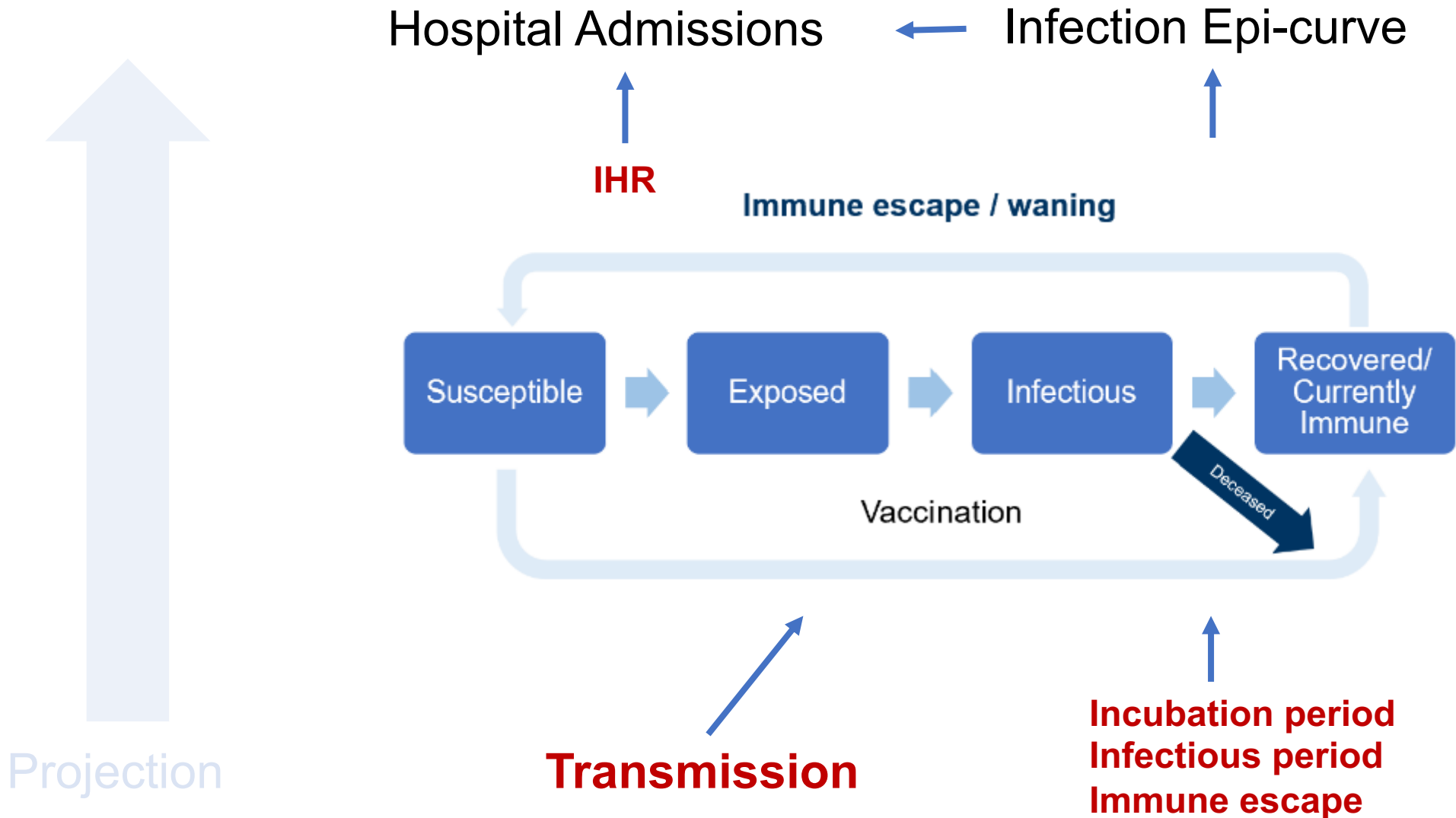
Immune escape / waning



Transmission

Incubation period
Infectious period
Immune escape

Case 2: Scenario projections



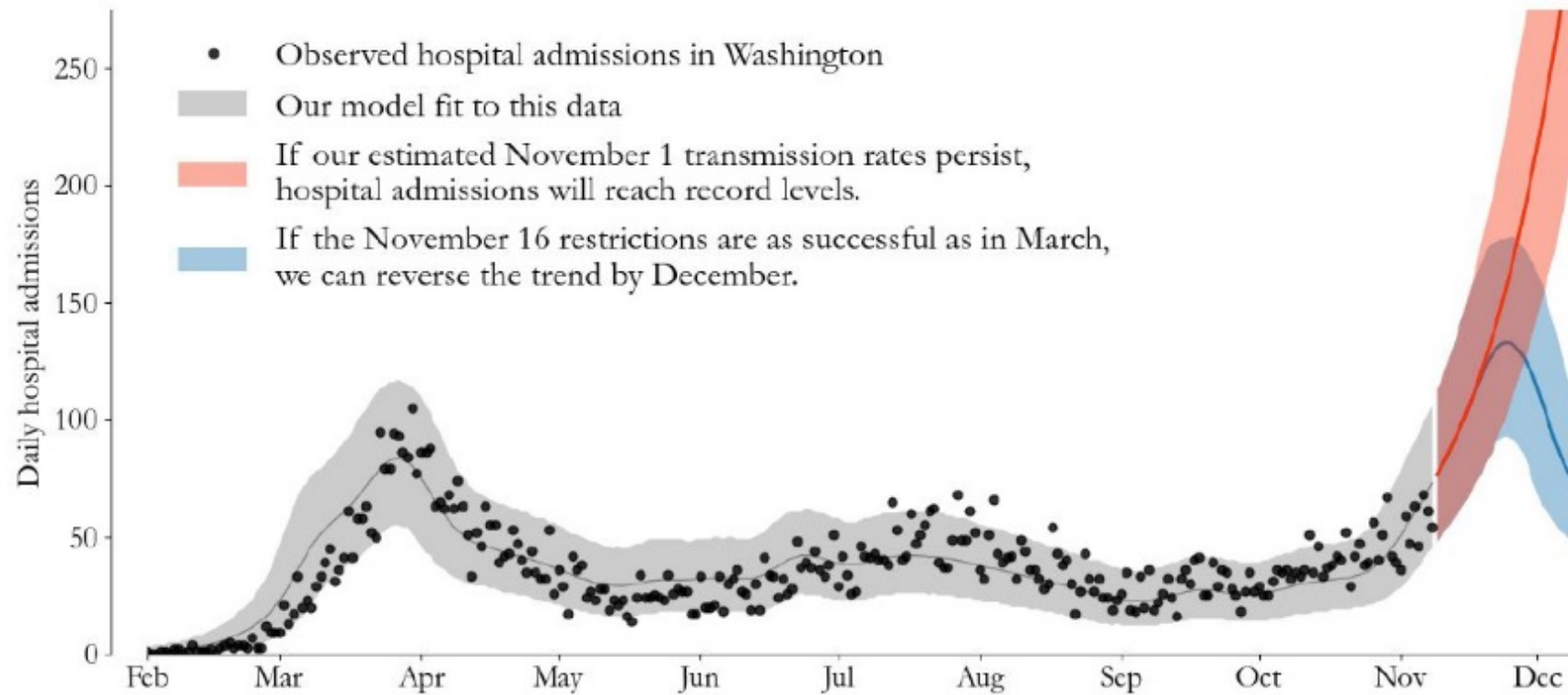
Case 2: Scenario projections

- Infections are projected directly from the model
- Hospital admissions are calculated based on future IHR
- Beds occupied are inferred from the hospital admissions and length of stay distribution, scaled to match WA health data



Case 1: Modeling to inform policy

November 16, 2020 restrictions

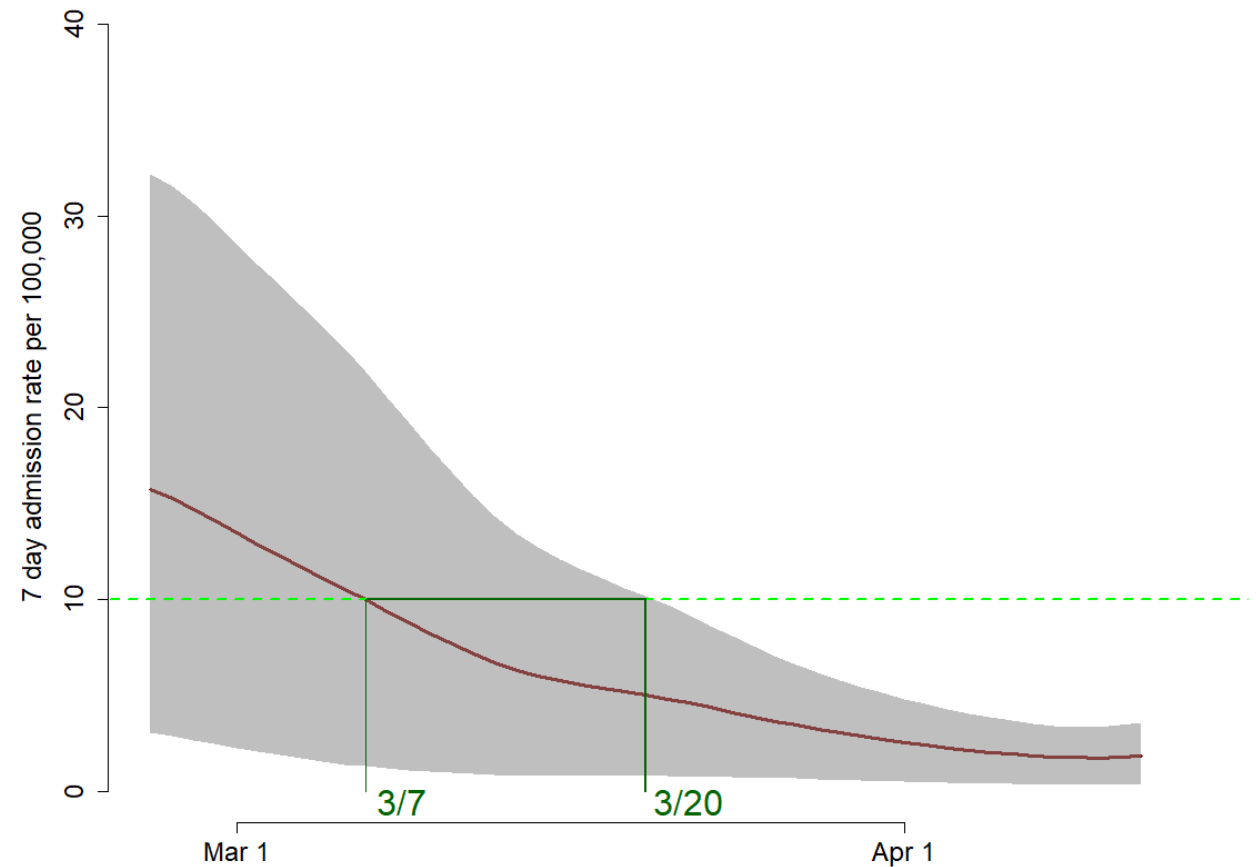


Case 2: Modeling to inform policy

Projected date at which admissions drop to CDC community criteria

- Model estimate: **10 admissions per 7 days per 100,000** between 2/25 – 3/20 (90% prediction interval) with a best estimate of **March 7**

Projected hospital admissions and 90% prediction interval





Question for the Audience

What is the biggest challenge your face in your work with data?

- A. Getting quality data
- B. Finding the right partnerships
- C. Having receptive leadership
- D. Insufficient technical support
- E. All of the above
- F. Other (Type in chat)

Equity: What do we *want* to know?

- Who is getting infected?
 - Age, race and ethnicity, employment, socioeconomic status
- Where is transmission happening?
 - Household, household size, workplace, type of work, other
- Who is being burdened by COVID-19?
 - Age, race and ethnicity, employment, socioeconomic status
- Who will be burdened going forward?
- How can we ensure equitable policy decisions?



Reflection

- Relationship with academic and not-for-profit modelers was very valuable to WADOH
- Importance of science communication
- Collaborations with other west coast states
- Benefit of receptive leadership
- **Challenges:** getting/giving data, data quality, model updates

Resources

Infectious Disease Models for Decision Makers (free online course)

<https://www.coursera.org/learn/infectious-disease-transmission-models-for-decision-makers>

Example modeling situation report for Washington State

<https://doh.wa.gov/sites/default/files/2022-03/820-114-SituationReport-20210922.pdf?uid=6480bd2762506>

University of Washington Summer Institute Statistics and Modeling for Infectious Diseases

<https://si.biostat.washington.edu/institutes/sismid>

General Infectious Disease Epidemiology and Modeling Resources

https://andreashandel.github.io/IDEMAcourse/General_Resources.html

QUESTIONS?



Q&A

To ask a question, please click the **Q&A icon** in the Zoom toolbar to open your Q&A Pod.