Applications of Nowcasting Methods to Notifiable Disease Surveillance

Ian Painter | Washington State Department of Health
Ben Stan, Yongzhe Wang, Fred Yu
Spring 2022

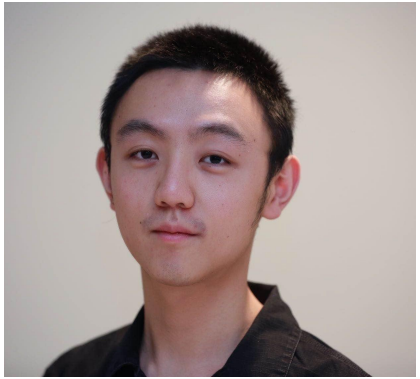
UNIVERSITY of WASHINGTON

Our team

Ian Painter

Ben Stan

Yongzhe Wang



Fred Yu

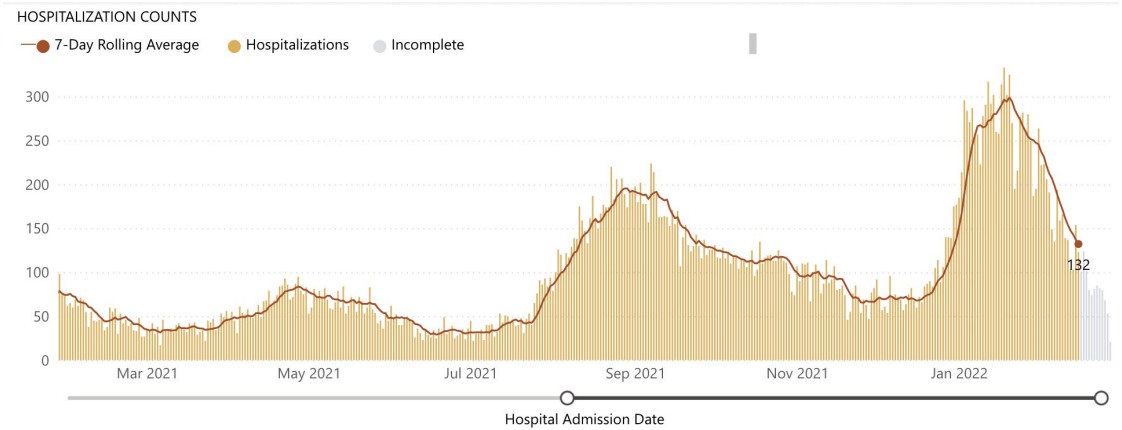
UNIVERSITY of WASHINGTON

Scientific Background

UNIVERSITY of WASHINGTON

Problem overview

- WA State DOH concerned with Covid-19 hospitalizations
  - Track by date of disease transmission to monitor disease spread
  - Positive test specimen collection date equated to disease transmission date for simplicity



Delay of interest

- Only concerned with cases that result in hospitalization
- Variable delay between test collection and hospitalization (1-14 days)

- Examples
  - Subject 1: Positive test on Day 1 → Hospital report on Day 10 → Delay = 9
  - Subject 2: Positive test on Day 9 → Hospital report on Day 11 → Delay = 2

Objectives

- Understand the true nature of the data and delays
- Develop nowcasting models to mitigate impact of delays
- Evaluate the performance and robustness of our models, and determine the highest-performing option
- Establish the contexts under which our model may be used by the DOH

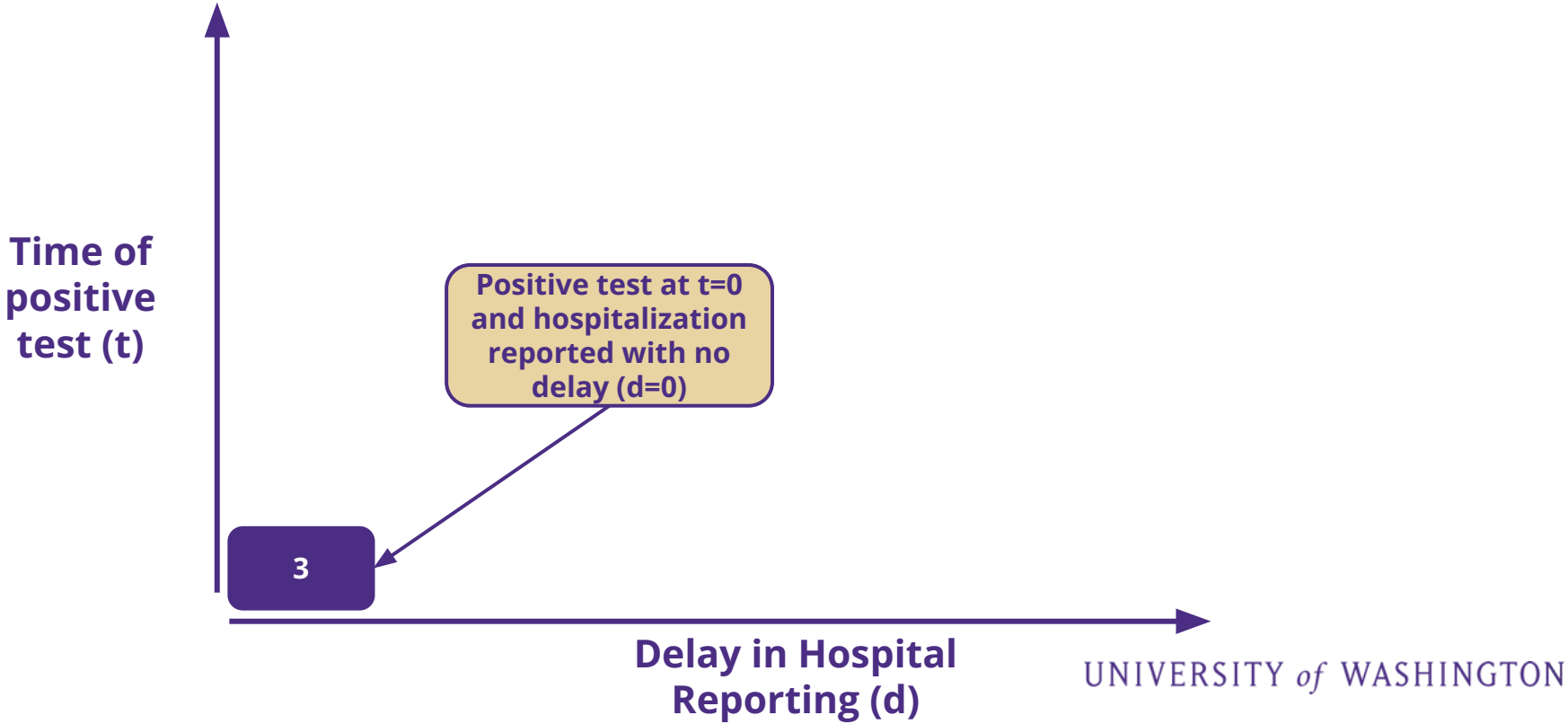
Nowcasting Background

UNIVERSITY of WASHINGTON

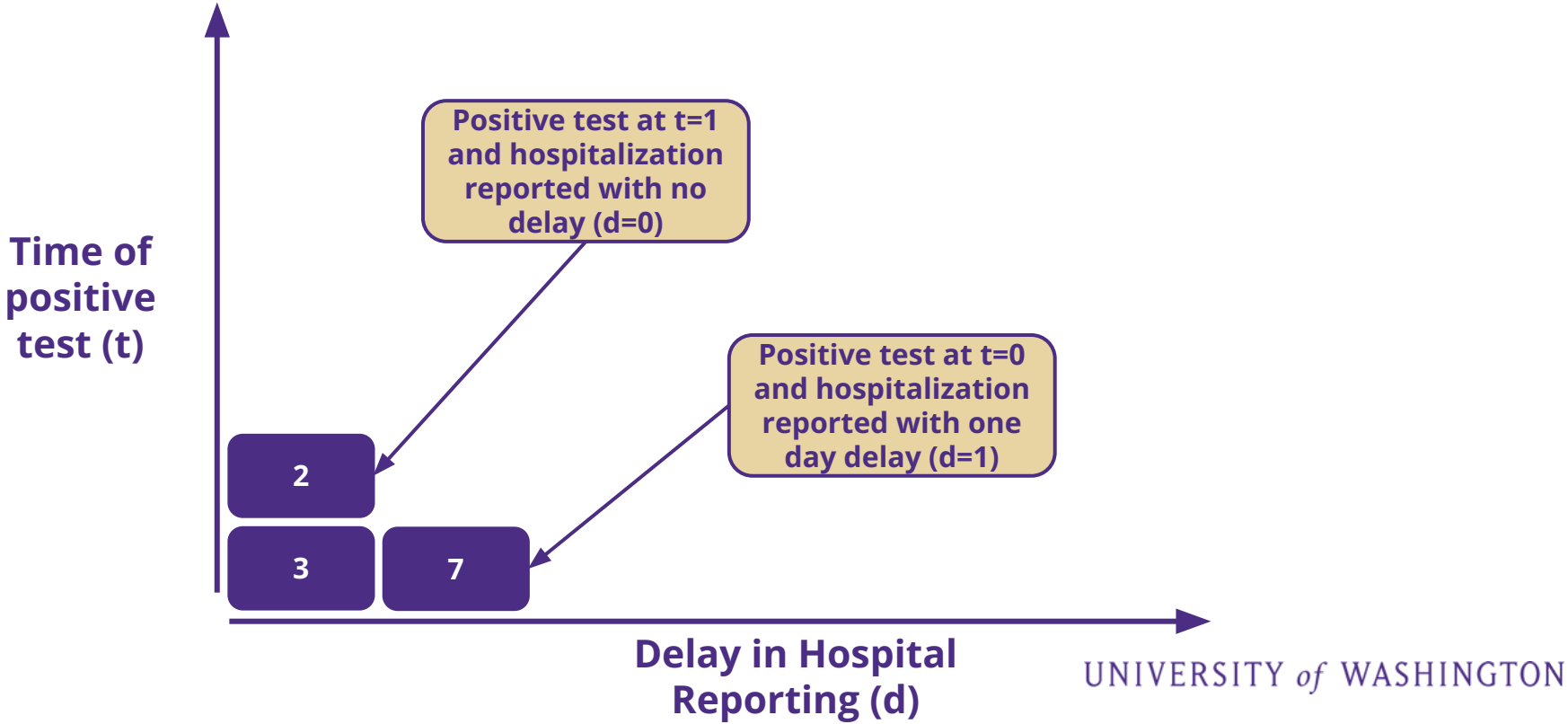
Nowcasting - Predicting the present

- Estimate the **occurred-but-not-yet reported** events
  - e.g. How many hospitalizations will result from the positive tests collected yesterday?
- Originated in actuarial literatures, found application when analysing AIDS/HIV epidemic
- Modeling the **delay distribution** is a common feature

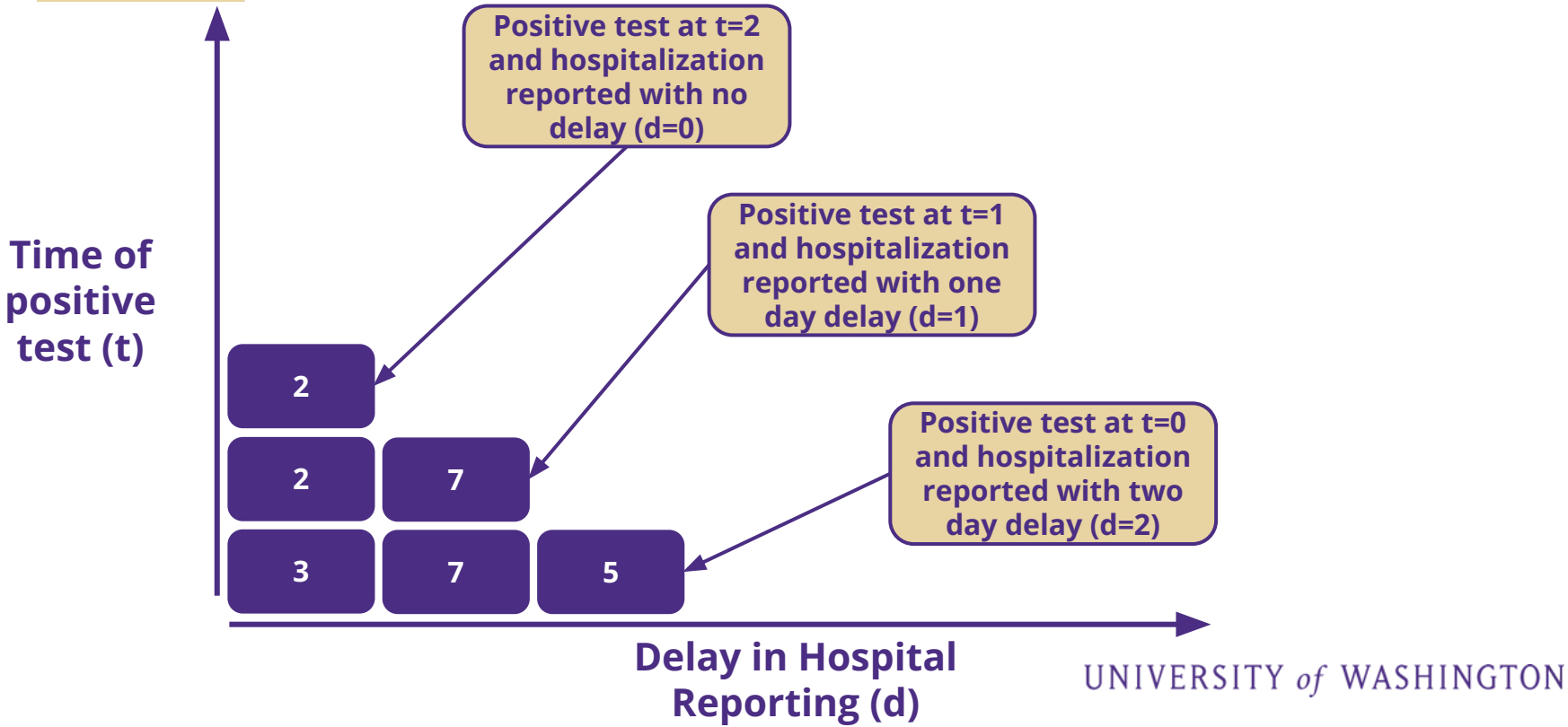
Nowcasting example – Day 1



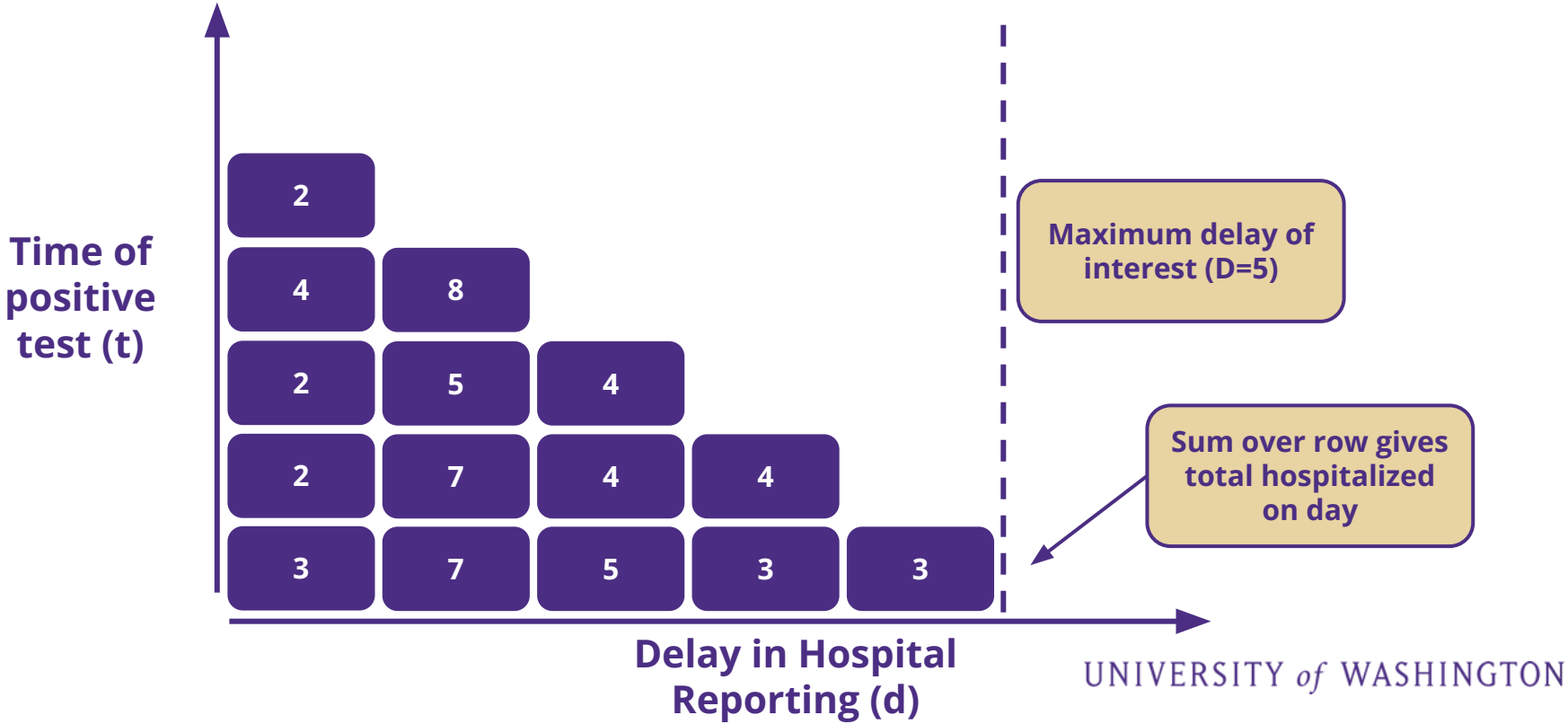
Nowcasting example – Day 2



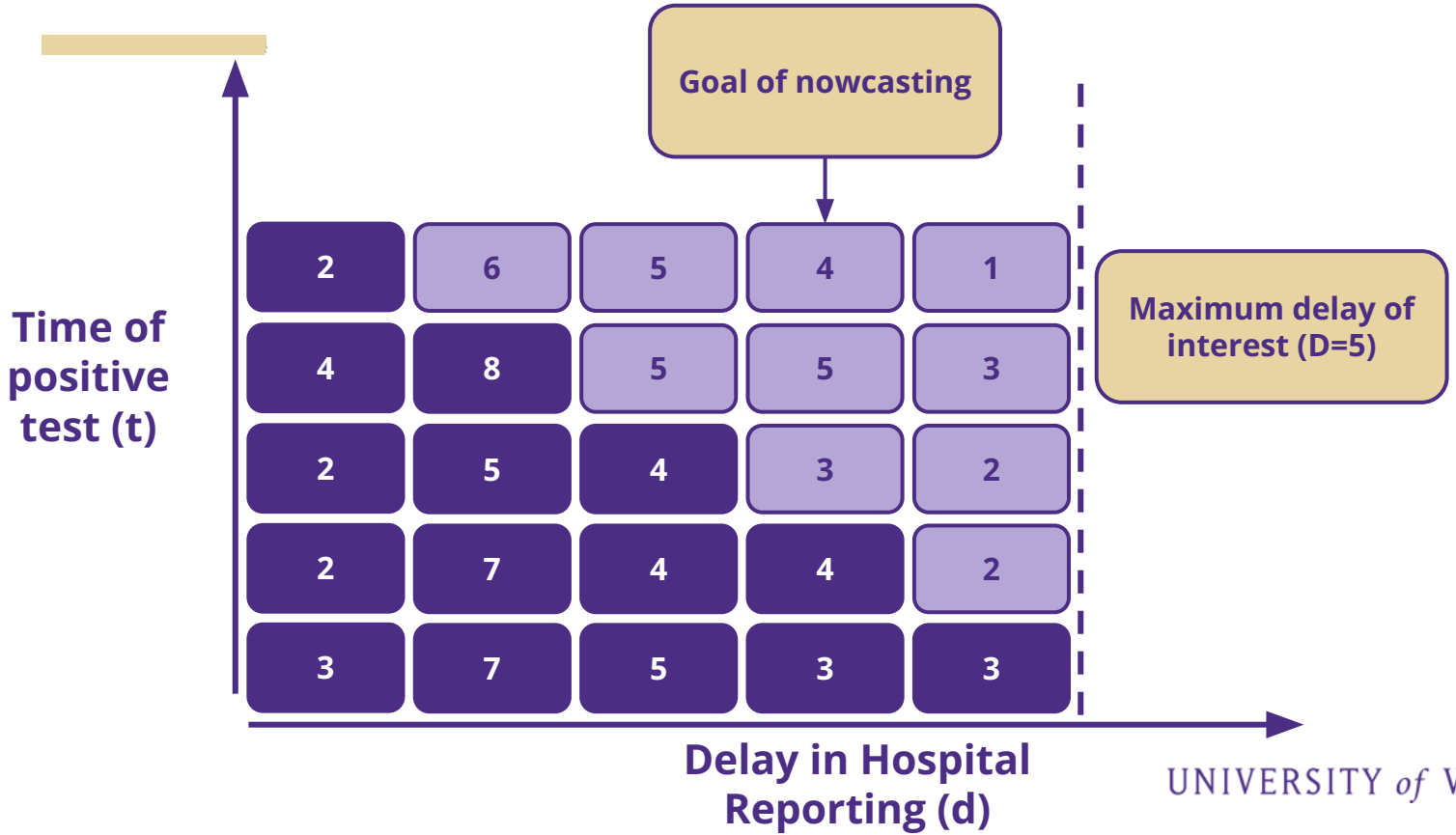
Nowcasting example – Day 3



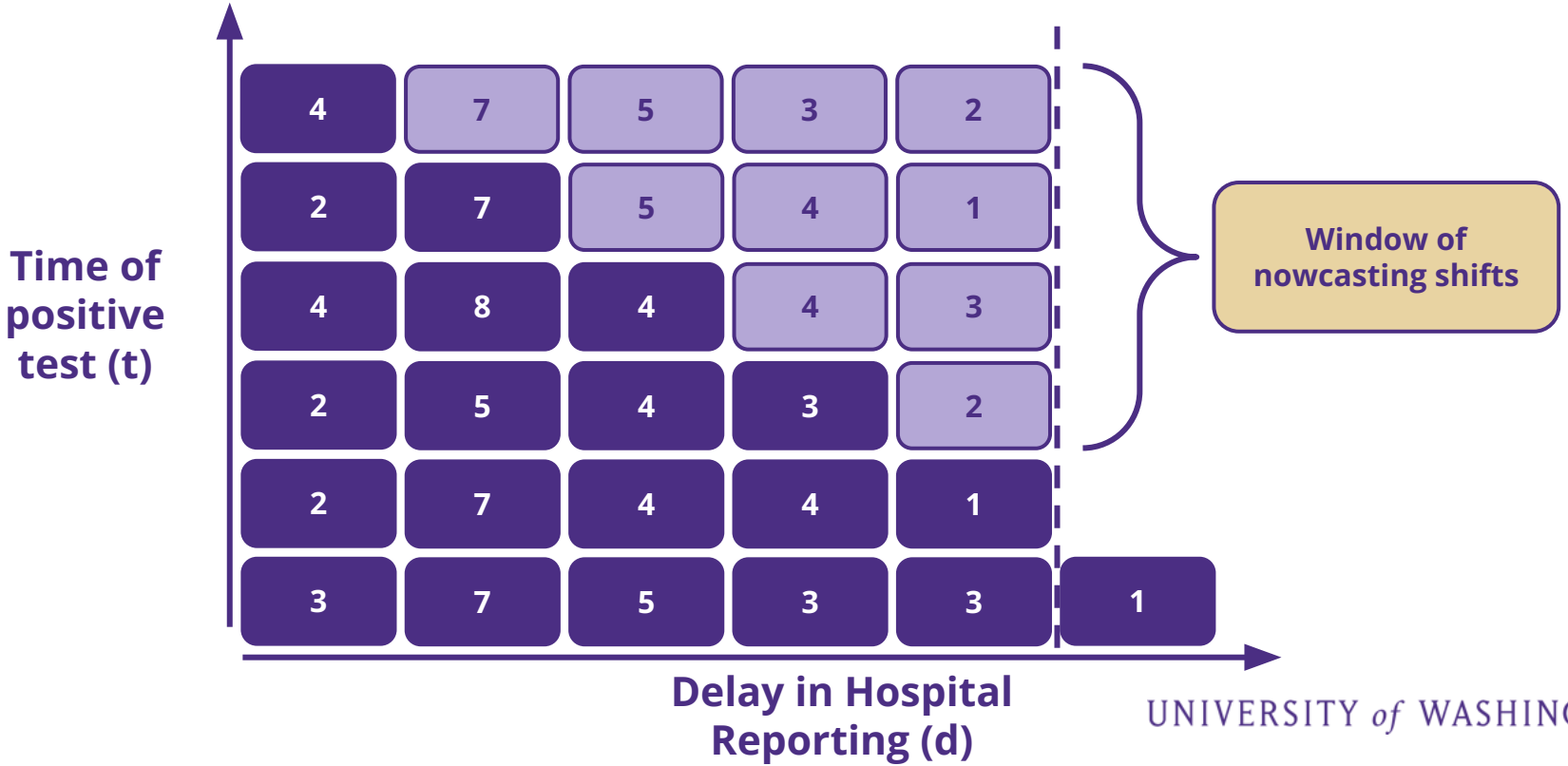
Nowcasting example – Day 5



Nowcasting example – Day 5



Nowcasting example – Day 6

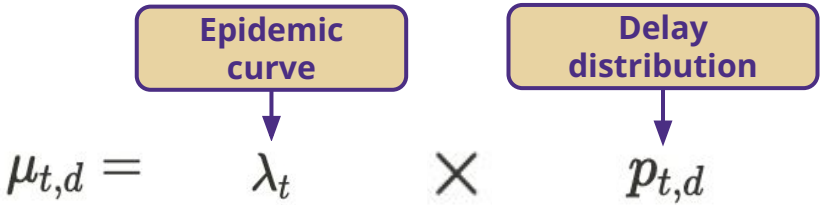


Nowcasting Methods

UNIVERSITY *of* WASHINGTON

General nowcasting framework

Assumption: *n(t, d)*, i.e. the case counts at day *t*, reported with *d* days of delay, follows an underlying in-homogeneous poisson process with the mean



Nowcasting can thus be done through the following steps:

Step 1: Estimate the rate of occurrence of cases at day *t*, i.e. the epidemic curve

Step 2: Estimate the proportions of cases at day t, delayed for d days, i.e. the delay distribution

Step 3: Predict the unobserved *n(t,d)*'s in order to compute the total *N(t, T+D)*

Bayesian nowcasting models

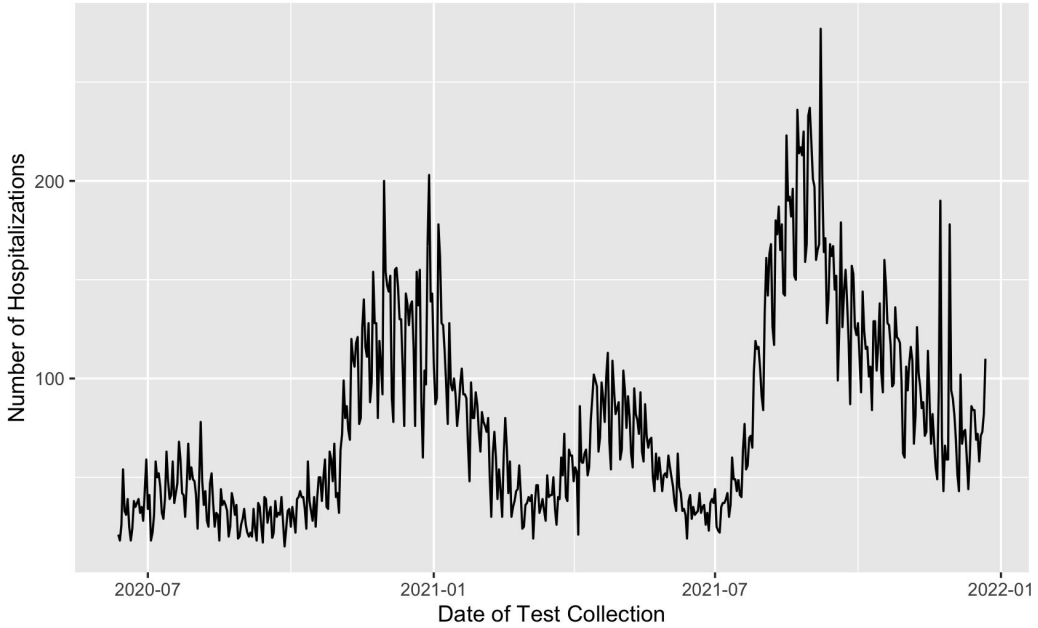
Model	Epidemic curve	Delay distribution
HH (*Hohle & Heiden*)	Assume a gamma mean $\lambda_t \sim Gamma(a_\lambda, b_\lambda)$ Then the true count given mean is Poisson.	- Assume **time-invariant** delay - Model using GD-Multinomial conjugacy framework
NobBS (Nowcasting by Bayesian Smoothing)	Model as a first-order random walk	- Assume **time-invariant** delay - Model using GD-Multinomial conjugacy framework

Results

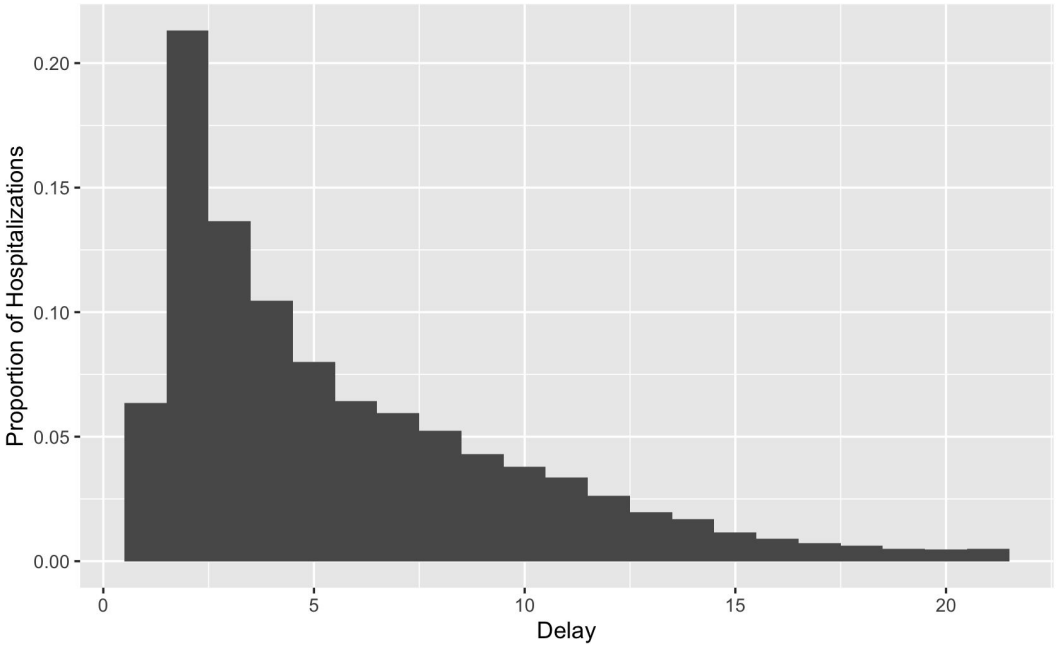
Data structure and exploratory analysis

- Schema: One row per hospitalization
- Time of interest: Test collection to hospitalization reporting
- Date range of data:

  6/11/20 to 1/12/22
- Max delay of interest: **21 days**
  - Analysis start: 7/2/20
  - Analysis end: 12/21/21

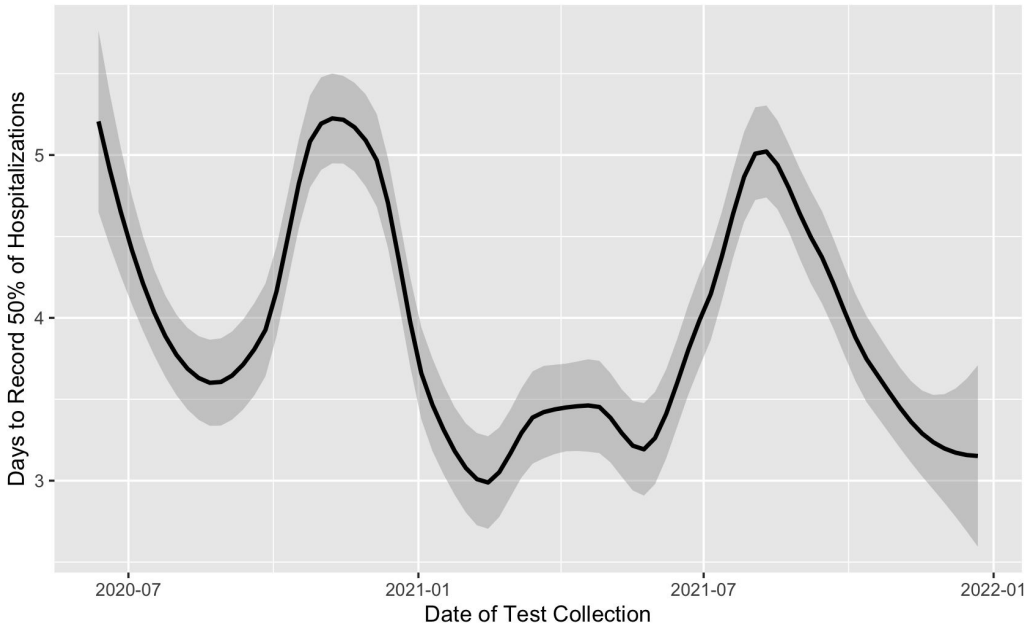


Exploratory analysis



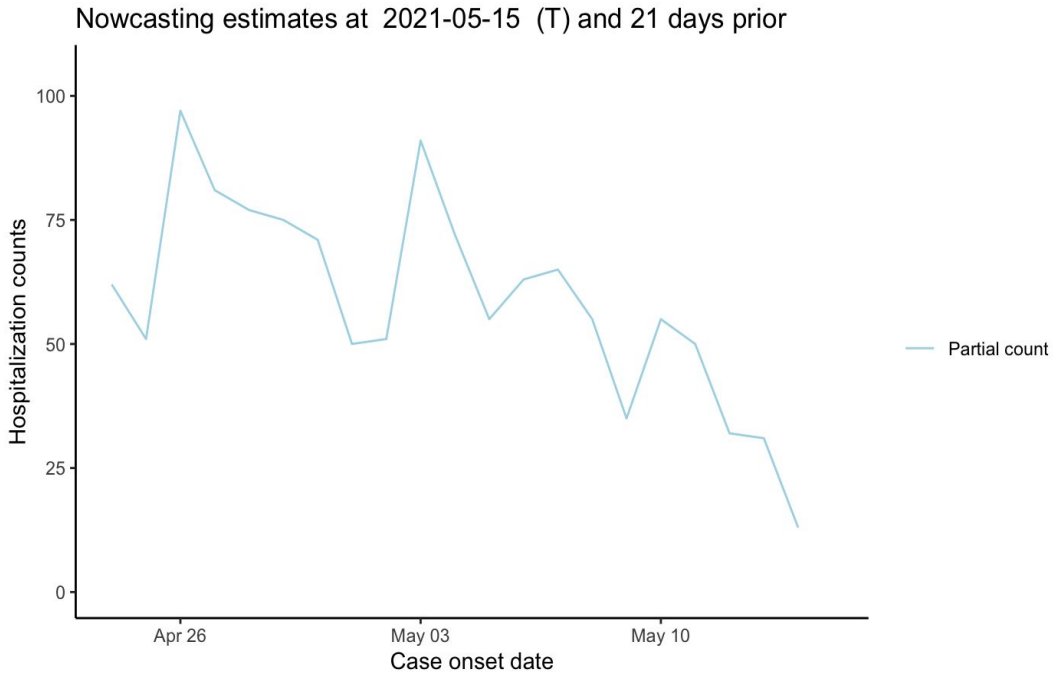
- No hospitalizations confirmed on date of test (delay = 0)
- Plurality of cases (20%+) reported with two days delay

Exploratory analysis



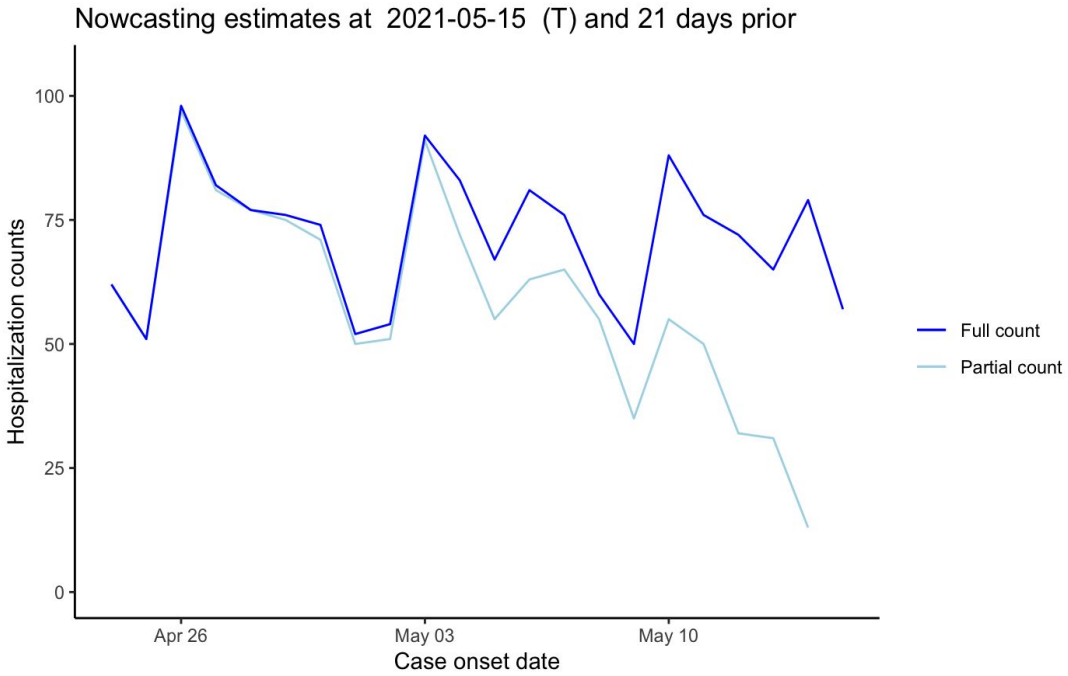
- Days needed to observe 50% of total hospitalizations fluctuates over pandemic
- Times associated with higher volumes require more days to reach 50% of hospitalizations

Example nowcast at May 15, 2021



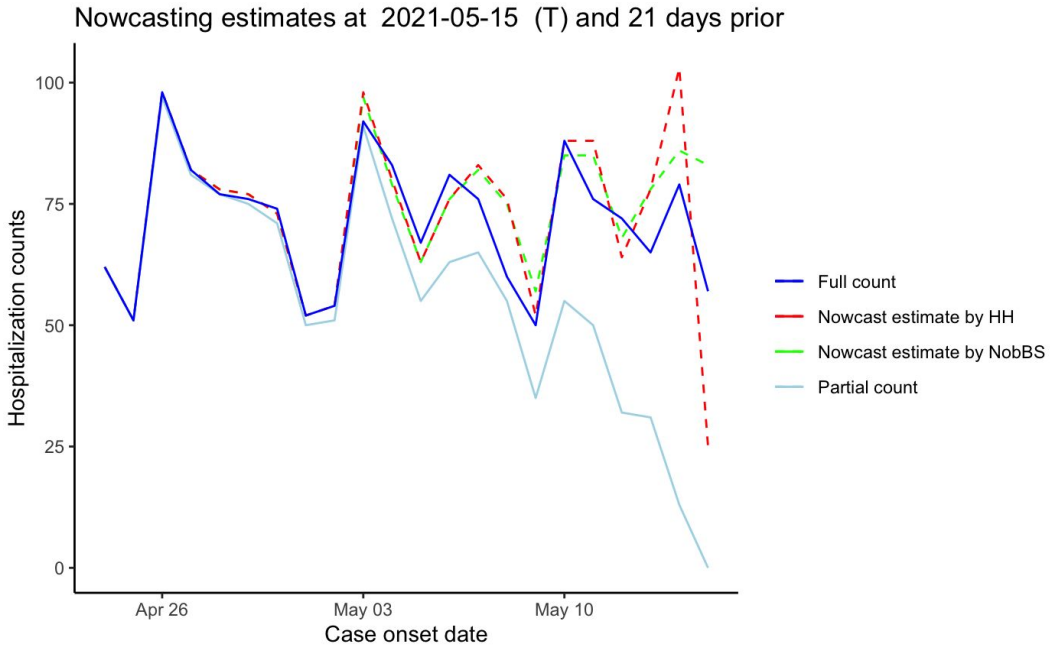
- Information available to us at time of nowcasting
- Decreases in most recent days

Example nowcast at May 15, 2021



- Lookback time: Days between date of nowcasting (e.g. 5/15) and date of prediction (ranges from 0 to 21)
- Gap between partial and full counts narrows as lookback time increases

Example nowcast at May 15, 2021



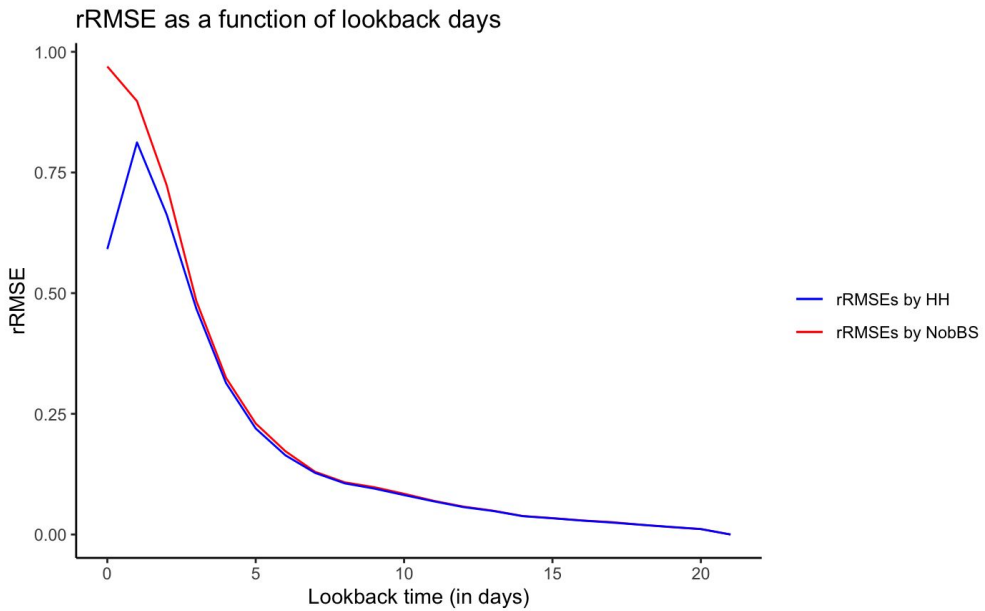
- Accuracy increases with increasing lookback time
- Direction of error (overpredict v underpredict) varies by date

Evaluation of nowcasts

- Given the max delay, **22** estimates generated for each date of nowcasting
- We use *relative root-mean-square error (**rRMSE**)* as the metric to evaluate the quality of different nowcasting models
  - Think of this as average relative error with penalty for large deviations
- The medians of *rRMSE* for all **539** times of nowcast

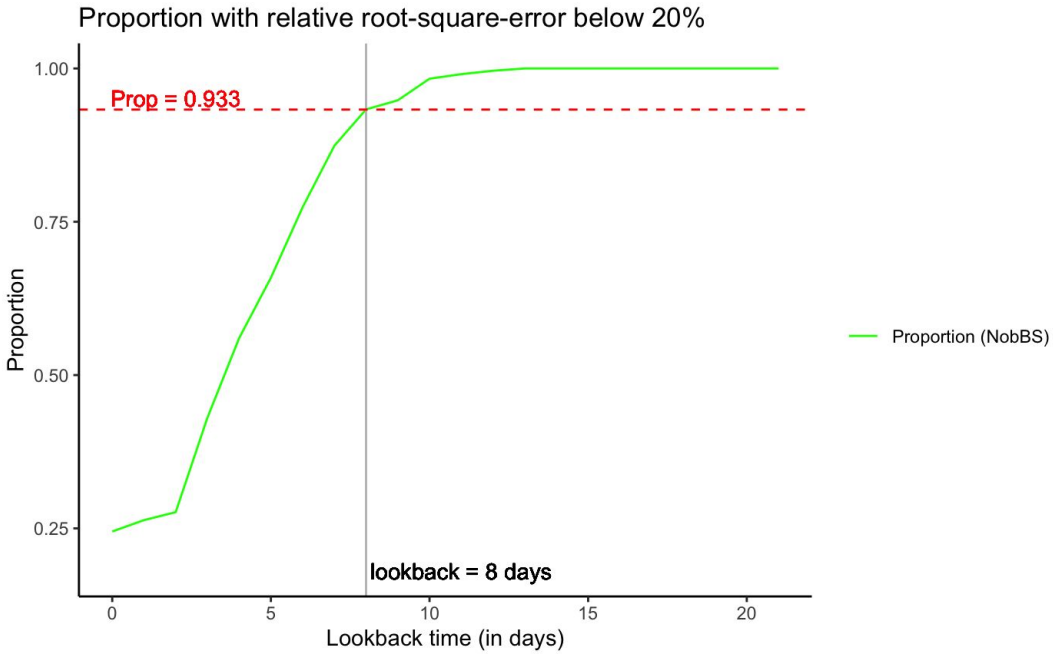
	HH	NobBS
rRMSE Median	0.261	0.197
rRMSE Median (lookback >= 5)	0.089	0.092

rRMSE by lookback time (in days)



- Both models see decrease in rRMSE as lookback time increases
- HH method appears more performant at shorter lookback times

Nowcasts with errors below threshold



- Collaborated with DOH to identify **20%** as tolerable error, as daily fluctuations reach this level
- For **lookback times of 8+**, we have high certainty of estimating within 20%
- Current DOH practice excludes most recent 16 days; **NobBS** model expedites decision-making by **8 days**

Next steps

- Handoff to WA State Department of Health
  - Determine parameters of use
  - Handoff code and guidelines
  - Productionize model and monitor
- Improve model
  - Incorporate reporting hospital information
  - Perform HH method with epidemic curve modeled by spline

Acknowledgement

Special Thanks to:

Ian Painter
Prof. Patrick Heagerty
Prof. Lloyd Mancl
Minh Vo

And our Capstone program cohort!

Questions?

UNIVERSITY of WASHINGTON

Appendix

UNIVERSITY of WASHINGTON