Dynamic Case-Control Sampling for Rapid Estimation of Vaccine Effectiveness Against an Emerging Infectious Disease Variant

Taylor M. Fortnam
Advised by: Joseph W. Hogan, Alyssa M. Bilinski
Thomas Trikalinos, Roberta DeVito, Laura C. Chambers
Additional Collaborators:
Ewa King, Richard C. Huard, Ellen Amore
Lisa M. Gargano, Jim McDonald, Philip A. Chan

March 26, 2025

Brown University School of Public Health
Department of Biostatistics

Table of Contents

Introduction

Background

- Analyzing public health surveillance data collected on COVID-19 comes with numerous challenges
- These challenges may impact the conclusions drawn through monitoring of COVID-19 indicators
- Analysis of COVID-19 data can inform current response efforts as well as future preparedness efforts

Background

- Analyzing public health surveillance data collected on COVID-19 comes with numerous challenges
- These challenges may impact the conclusions drawn through monitoring of COVID-19 indicators
- Analysis of COVID-19 data can inform current response efforts as well as future preparedness efforts

We applied statistical approaches to address pertinent questions formed through ongoing conversations with the Rhode Island Department of Health over the course of the pandemic using real-world COVID-19 data.

Motivating Question

Can genomic surveillance data be used to generate real-time updates of vaccine effectiveness against an emerging variant?

Objective

Dynamic Case-Control Sampling for Rapid Estimation of Vaccine Effectiveness Against an Emerging Infectious Disease Variant

Objective:

Use case-control sampling to produce dynamically updating estimates of the effectiveness of COVID-19 vaccines against infection with an emerging SARS-CoV-2 variant in Rhode Island and compare to vaccine effectiveness against previous variants

Active Monitoring

Surveillance Data

Statistical Modeling

Rhode Island Focus

Background

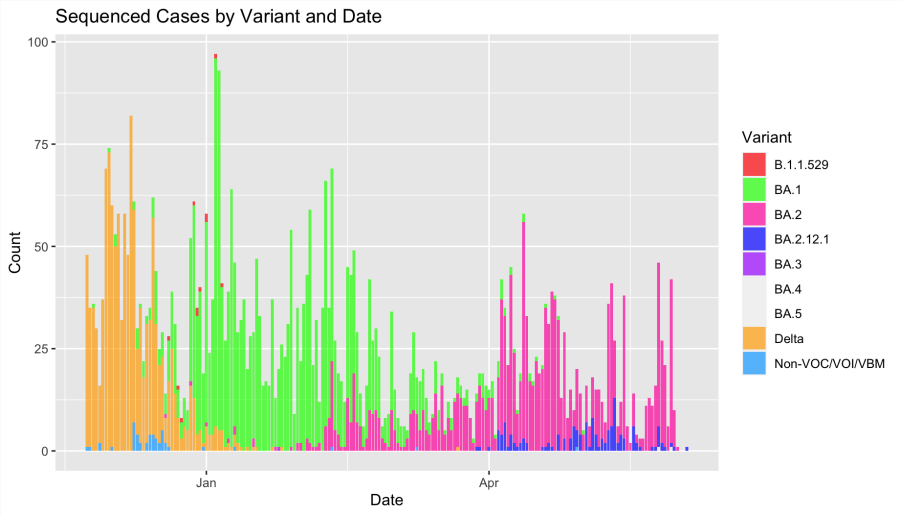
- New COVID-19 variants arise frequently with different viral properties that can impact the effectiveness of existing vaccines
- Public health officials must rapidly assess vaccine effectiveness (VE) against new variants so that they can adjust mitigation measures
- We propose a dynamically-updating method to use genomic surveillance data to produce estimates of VE against any infection with an emerging variant
- We apply this method to the BA.1 and BA.2 sub-lineages of the Omicron variant

Motivation

Limitations of traditional methods for estimating VE:

- Obtaining reliable estimates of VE often involves conducting a prospective cohort or test-negative case-control study, both of which require large sample sizes and substantial time for cases to accumulate
- Genomic sequencing is costly and typically only available for a subsample of positive cases

Background



Cases identified as particular variants and sub-lineages over time.

Key Idea

In the context of an emerging variant:

- Have reliable information about VE against a previous-circulating (index) variant from previous studies
- Want timely estimates of vaccine effectiveness against the emerging variant

We can estimate VE for an emerging variant relative to the index variant

Methods

Data

Surveillance Data from RIDOH:

- SARS-CoV-2 positive specimens linked with vaccination registry
- Associated demographic information: age, sex, race, congregate care status, and zip-code based community risk classification

Description of Samples:

- Only utilize first diagnosed infections in analysis
- 5,751 individuals ages 16 and over with a sequenced sample for their first diagnosed infection reported to RIDOH
  - 2,220 (39%) BA.1 sub-lineage
  - 1,462 (25%) BA.2 sub-lineage
  - 2,069 (36%) Delta variant

Study Design and Notation

S denotes variant (or subtype) of sequenced virus from a person infected with SARS-CoV-2

- $S = 0$ corresponds to uninfected
- $S = s$ corresponds to infection with an index variant
- $S = s'$ corresponds to infection with an emerging variant

V denotes a nominal categorical variable representing vaccine status

- $V \in \{0, 1, 2, \ldots, J\}$ represents level of vaccination received
- $V = 0$ corresponds to unvaccinated
- In our application, $V = 2$ represents full vaccination

Study Design and Notation

VE against subtype s:

$$\begin{aligned} \mathrm{VE}_v(s) &= 1 - \frac{P(S = s|V = v)}{P(S = s|V = 0)} \\ &= 1 - \mathrm{RR}_v(s, 0) \end{aligned}$$

$\mathrm{RR}_v(s, 0)$ indicates that VE is calculated in terms of risk of infection with subtype s relative to not being infected.

When risk of infection is low, VE can be expressed in terms of an odds ratio,

$$\begin{aligned} \mathrm{VE}_v(s) &= 1 - \frac{P(S = s \mid V = v)/P(S = 0 \mid V = v)}{P(S = s \mid V = 0)/P(S = 0 \mid V = 0)} \\ &= 1 - \psi_v(s, 0) \end{aligned}$$

Estimating Vaccine Effectiveness

Objective: estimate VE against infection with an emerging variant s':

$$\mathrm{VE}_v(s') = 1 - \psi_v(s', 0)$$

where

$$\begin{aligned}
\psi_v(s', 0) &= \frac{P(S = s' \mid V = v)/P(S = 0 \mid V = v)}{P(S = s' \mid V = 0)/P(S = 0 \mid V = 0)} \\
&= \frac{P(S = s' \mid V = v)/P(S = s \mid V = v)}{P(S = s' \mid V = 0)/P(S = s \mid V = 0)} \times \frac{P(S = s \mid V = v)/P(S = 0 \mid V = v)}{P(S = s \mid V = 0)/P(S = 0 \mid V = 0)} \\
&= \psi_v(s', s)\psi_v(s, 0)
\end{aligned}$$

Then,

$$\begin{aligned} \mathrm{VE}_v(s') &= 1 - \psi_v(s', s)\psi_v(s, 0) \\ &= 1 - \psi_v(s', s)\left\{1 - \mathrm{VE}_v(s)\right\} \end{aligned}$$

Then,

$$\begin{aligned} \mathrm{VE}_v(s') &= 1 - \psi_v(s', s)\psi_v(s, 0) \\ &= 1 - \psi_v(s', s)\left\{1 - \mathrm{VE}_v(s)\right\} \end{aligned}$$

To estimate $\mathrm{VE}_v(s')$:

- Accumulating data need to contain samples of both s' and s
- Need to be able to find estimates of $\mathrm{VE}_v(s)$ in the literature, usually estimated on a different population

Considerations when Estimating $VE_v(s')$

- **Motivation:** Not enough information to estimate $\psi(s', 0)$ in the early phase of emerging variant
- **Approach:** Dynamic matched case-control analysis to update estimates of $\psi_v(s', s)$ as infections from emerging variant accumulate
  - Cases: Emerging variant (Omicron)
  - Controls: Index variant (Delta)
- Subset of those with sequenced virus potentially nonrandom relative to the population of interest
- Uncertainty comes from two sources:
  1. Uncertainty in estimate of VE against index variant s
  2. Uncertainty associated with $\psi_v(s', s)$

Analysis Strategy Overview

Address nonrandom selection of those with sequences

- Use inverse probability weighting applied to entire sample of infections

Matching cases and controls

- Use full optimal matching based on propensity scores
- Ensures all sequenced observations are used
- Matched sets may be unbalanced

Estimate $\psi_v(s', s)$

- Use weighted logistic regression
- Weighting for sample selection and unbalanced matched sets

Producing an Estimate of $\psi_v(s', s)$

Weighted logistic regression on the matched dataset:

$$\begin{aligned} \text{logit}\{P(S = s' \mid X_i, V_i)\} &= \sum_{v=0}^{J} \theta_v \mathbb{I}(V_i = v) + h(\boldsymbol{X}_i; \boldsymbol{\alpha}) \\ &= \sum_{v=0}^{J} \log(\psi_v(s', s)) \mathbb{I}(V_i = v) + h(\boldsymbol{X}_i; \boldsymbol{\alpha}) \end{aligned}$$

where $h(X_i; \alpha)$ indicates the function for covariate adjustment

Producing an Estimate of $\psi_v(s', s)$

Weighted logistic regression on the matched dataset:

$$\begin{aligned} \text{logit}\{P(S = s' \mid X_i, V_i)\} &= \sum_{v=0}^{J} \theta_v \mathbb{I}(V_i = v) + h(\boldsymbol{X}_i; \boldsymbol{\alpha}) \\ &= \sum_{v=0}^{J} \log(\psi_v(s', s)) \mathbb{I}(V_i = v) + h(\boldsymbol{X}_i; \boldsymbol{\alpha}) \end{aligned}$$

where $h(X_i; \alpha)$ indicates the function for covariate adjustment

Exponentiated coefficients from this model, $\hat{\psi}_v(s', s)$, provide a measure of VE against infection with the index variant relative to the emerging variant

Estimating Uncertainty Associated with $\text{VE}_v(s')$

Approximate sampling distributions using fitted model and published studies:

- Normal distribution for $\log\left\{\hat{\psi}_v(s', s)\right\}$: $\mathcal{N}(\hat{\theta}_v, \hat{\sigma}^2_{\theta_v})$, where $\sigma^2_{\hat{\theta}_v} = \text{var}(\hat{\theta}_v)$
- Normal distribution for $\widehat{\mu}_v = \log\left\{\widehat{\text{VE}}_v(s)\right\}$: $\mathcal{N}(\hat{\mu}_v, \hat{\sigma}^2_{\mu_v})$

Drawing 95% CI for $\text{VE}_v(s')$:

1. Simulate a pair $(\tilde{\theta}_v, \tilde{\mu}_v)$
2. Set $\tilde{\psi}_v(s', s) = \exp(\tilde{\theta}_v)$ and $\widetilde{\text{VE}}_v(s) = \exp(\tilde{\mu}_v)$
3. Calculate $\widetilde{\text{VE}}_v(s')$ from $\tilde{\psi}_v(s', s)$ and $\widetilde{\text{VE}}_v(s)$
4. Compute $.025^{th}$ and $.975^{th}$ quantiles

Results

Estimated Odds Ratios, $\psi_v(s', s)$

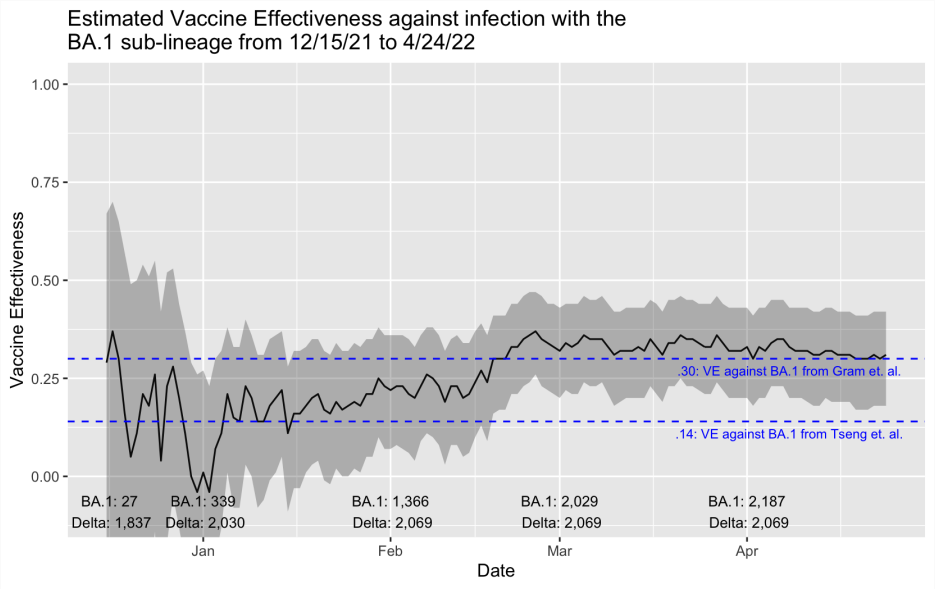
$\psi(s', s) > 1$ indicates that VE against the Delta variant is greater than against the Omicron variant.

	Estimate (CI) for $\psi(s', s)$	
Vaccination Status	$s' =$ BA.1	$s' =$ BA.2
Unvaccinated	–	–
One Dose of Two-Dose Series	3.77 (2.72, 5.27)	5.13 (3.49, 7.58)
Completed Primary Series	1.90 (1.64, 2.20)	1.24 (1.02, 1.51)

Estimated Vaccine Effectiveness

Location	Study Type	# infections; # without infection	Vaccine effectiveness		
			Delta ($VE_v(s)$)	BA.1 ($VE_v(s')$)	BA.2 ($VE_v(s')$)
California [1]	cohort, Delta-dominant	197,535; 2,919,754	49 (46, 51)	3 (–13, 17)	37 (23, 48)
California [2]	case control, sequenced	2,027; 10,135	87 (84, 89)	75 (69, 81)	84 (79, 88)
California [3]	case-control, S-Gene Target Failure	26,683; 109,662	64 (60, 67)	31 (18, 42)	55 (44, 64)
Minnesota [4]	case control Delta-dominant	25,869; 25,869	59 (36, 75)	20 (–28, 52)	47 (15, 69)
Norway [5]	cohort, sequenced	5,430 4,199,429	65 (61, 68)	33 (21, 44)	56 (46, 65)
Denmark [6]	cohort; Delta-dominant and sequenced	34,636 842,397	65 (64, 66)	33 (23, 43)	56 (47, 64)

Progression of Estimate of $VE_v(s')$ as Data Accumulate



Discussion

Discussion

- Can produce estimates of VE that stabilize quickly and are comparable in magnitude to results produced by other methods
- Able to detect reduced VE against each of the BA.1 and BA.2 sub-lineages relative to the Delta variant
- Our estimates have large associated error, this could be reduced by sequencing a higher proportion of cases or implementing the method in a larger health department with access to more case records

Limitations

- The precision of our estimate depends on the precision of estimates reported in the literature
- We have assumed that estimates of VE against the index variant are transportable to the Rhode Island population and that the vaccine effect is durable
- Sequencing delays can be substantial

Future Work

- Address transportability issue
- Determine properties under larger sample sizes
- Evaluate how well the method does under assumption violations

Conclusion

Conclusion

How to address an active surveillance question in real-time?

- COVID-19 data are extensive and present many challenges
- We can address these with statistical and mathematical approaches in order to still make appropriate use of the data
- It is important to analyze these data to evaluate novel policy approaches help inform future health policy responses
- Analysis approaches need to consider information pertinent to public health responses and modeling strategies need to be adapted to account for data limitations

References

References

[1] Sara Y Tartof, Laura Puzniak, Vennis Hong, Timothy B. Frankland, Bradley K. Ackerson, Harpreet S. Takhar, Oluwaseye A. Ogun, Sarah R. Simmons, Joann M. Zamparo, Sharon Gray, Srinivas R. Valluri, Kaije Pan, Luis Jodar, and John M. McLaughlin. Effectiveness of a third dose of BNT162b2 mRNA COVID-19 vaccine in a large US health system: A retrospective cohort study. *The Lancet Regional Health –Americas*, 2022/03/08 . doi: 10.1016/j.lana.2022.100198. URL https://doi.org/10.1016/j.lana.2022.100198.

[2] Katia J Bruxvoort, Lina S Sy, Lei Qian, Bradley K Ackerson, Yi Luo, Gina S Lee, Yun Tian, Ana Florea, Michael Aragones, Julia E Tubert, Harpreet S Takhar, Jennifer H Ku, Yamuna D Paila, Carla A Talarico, and Hung Fu Tseng. Effectiveness of mRNA-1273 against Delta, Mu, and other emerging variants of SARS-CoV-2: test negative case-control study. *BMJ*, 375, 2021. doi: 10.1136/bmj-2021-068848. URL `https://www.bmj.com/content/375/bmj-2021-068848`.

[3] Hung Fu Tseng, Bradley K. Ackerson, Yi Luo, Lina S. Sy, Carla A. Talarico, Yun Tian, Katia J. Bruxvoort, Julia E. Tubert, Ana Florea, Jennifer H. Ku, Gina S. Lee, Soon Kyu Choi, Harpreet S. Takhar, Michael Aragones, and Lei Qian. Effectiveness of mRNA-1273 against SARS-CoV-2 Omicron and Delta variants. *Nature Medicine*, 28(5):1063–1071, 2022. doi: 10.1038/s41591-022-01753-y. URL `https://doi.org/10.1038/s41591-022-01753-y`.

References iii

[4] Arjun Puranik, Patrick J. Lenehan, Eli Silvert, Michiel J.M. Niesen, Juan Corchado-Garcia, John C. O'Horo, Abinash Virk, Melanie D. Swift, John Halamka, Andrew D. Badley, A.J. Venkatakrishnan, and Venky Soundararajan. Comparison of two highly-effective mRNA vaccines for COVID-19 during periods of Alpha and Delta variant prevalence. *medRxiv, 2021. URL https://www.medrxiv.org/content/early/2021/08/21/2021.08.06.21261707.*

[5] Elina Seppälä, Lamprini Veneti, Jostein Starrfelt, Anders Skyrud Danielsen, Karoline Bragstad, Olav Hungnes, Arne Michael Taxt, Sara Viksmoen Watle, and Hinta Meijerink. Vaccine effectiveness against infection with the Delta (B.1.617.2) variant, Norway, April to August 2021. Eurosurveillance, 26(35):2100793, 2021. doi: https://doi.org/10.2807/1560-7917.ES.2021.26.35.2100793. URL https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.35.2100793.

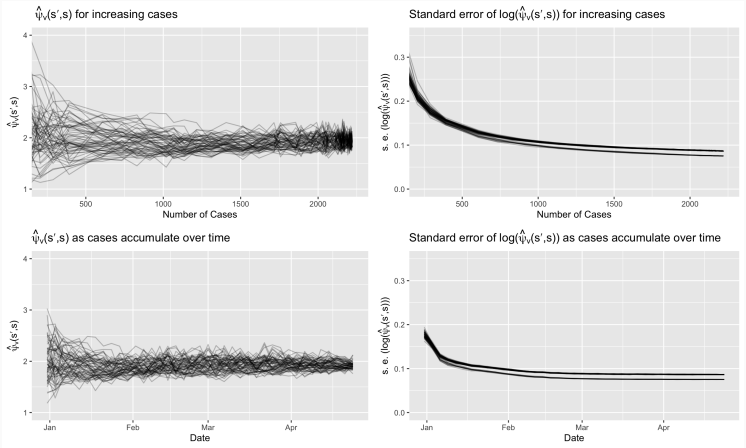
[6] *Mie Agermose Gram, Hanne-Dorthe Emborg, Astrid Blicher Schelde, Nikolaj Ulrik Friis, Katrine Finderup Nielsen, Ida Rask Moustsen-Helms, Rebecca Legarth, Janni Uyen Hoa Lam, Manon Chaine, Aisha Zahoor Malik, Morten Rasmussen, Jannik Fonager, Raphael Niklaus Sieber, Marc Stegger, Steen Ethelberg, Palle Valentiner-Branth, and Christian Holm Hansen. Vaccine effectiveness against SARS-CoV-2 infection or COVID-19 hospitalization with the Alpha, Delta, or Omicron SARS-CoV-2 variant: A nationwide Danish cohort study. PLoS Med, 19 (9):e1003992, Sep 2022. ISSN 1549-1676 (Electronic); 1549-1277 (Print); 1549-1277 (Linking). doi:* $10.1371/journal.pmed.1003992$.

Acknowledgments

Thank you!

Additional Slides

Simulation Study



Under randomly-shuffled case accumulation, $\psi_v(s', s)$ stabilizes around 2 and the standard error of $\psi_v(s', s)$ decreases as cases accumulate.