How Important Are Study Designs?

A Simulation Approach in Assessing Bias in VE Estimation with Time-Varying Vaccine Coverage, and Heterogeneous Testing and Baseline Attack Rates

Jing Lian Suah[a] Naor Bar-Zeev[b] Masliyana Husin[a]
Boon Hwa Tng[a] Maria Deloria Knoll[b]† Sheamini Sivasampu[a]†

[a]Institute for Clinical Research, National Institutes of Health, Malaysia
[b]Bloomberg School of Public Health, John Hopkins University
† Co-Senior Authors

June 27, 2022

Any views expressed are that of the researchers, and should not be taken to represent those of the Government of Malaysia

Highlights

- This paper simulated a **theoretical model with frictions in vaccination, testing, baseline disease risks, and heterogeneous vaccine effectiveness** to evaluate estimation bias across **four cohort and two test-negative designs**.
- In theory, **bias depends on behavioural asymmetries** (in testing, and baseline risk) between the vax-willing and vax-unwilling, and the **speed of vaccination rollout**.
- Even under 'ideal' conditions (no frictions), there is **intrinsic estimation bias** across all study designs of different magnitudes.
- In scenarios that may be reflective of past SARS-CoV-2 waves, the **degree of bias can be substantial**, attributable to variation in assumed testing and baseline risk frictions.
- A formal regression-based decomposition indicates that study designs have visibly **different primary sources of estimation bias, and degree of robustness** in general.
- This study warrants a **re-benchmarking of methodology and reporting checklists** for VE research, and informs the design of cost-effective surveillance by **quantifying part of the bias-implementation cost trade-off**.

Research Question and Approach

Research Question

How do commonly used VE study designs (variations of cohorts, and TNDs) perform in the presence of (i) time-varying vaccine coverage, (ii) latent selection of testing and risk behaviours on willingness to vaccinate, and (iii) 'leaky' vaccines?

Analytical Steps

1. Deriving an expression for the theoretical bias in a conceptual model of the epidemiological environment
2. Simulate data using the conceptual model, and compute the VE estimation bias under various configurations of cohorts and TNDs
3. Analyse the relative importance of the various drivers of bias

Logic of the entire study

k parameter sets $\Xi = (\Xi_1\ \Xi_2\ \Xi_3\ ...\ \Xi_k)$, as **input**, which contains the mean of the VE distribution, μ_{VE}, which is used to calculate estimation bias ($\widehat{VE} - \mu_{VE}$)

$\Downarrow$

Set up a **theoretical model** detailing the 'epidemiological environment' as a function $f(.)$, which describes how people get vaccinated, get tested, develop outcomes (generalisable), and how vaccines confer protection at the individual- & population-levels, etc

$\Downarrow$

Use the theoretical model f to **simulate data** from every parameter set in Ξ, producing k sets of simulated data (y_k); use **parallel processing** to speed up computation

$\Downarrow$

For every set of simulated data y_k, **estimate VE** using M study designs (variations of cohorts and TNDs), and **calculate the respective estimation bias**

$\Downarrow$

Concatenate all of the estimation biases by study designs, and parameter sets for **specific analysis**

$\Downarrow$	$\Downarrow$	$\Downarrow$
Scenario analysis: Zoom into special parameter sets e.g., similar to SARS-CoV-2 waves	**Drivers decomposition:** Linear regression to analyse relative importance of sources of bias	**Relative ranks (Additional):** Compute stability of relative bias between study designs

Computational logistics of the study

- The simulation was run on a AWS EC2 c6i.32xLarge Amazon Linux instance (128 vCPUs), which took **415516 seconds (115.42 hours)**.
  - Running on a Windows 10, 4-core (8 logical processors) 11th gen i5 local machine would have taken **at least 1154 hours (48 days)**, subject to sufficient memory.
  - Virtual machines can be expensive! This means that stress tests, fail-safes, and minimum working examples need to be implemented thoroughly beforehand.
- Post-simulation analyses were run on Windows 10, 4-core (8 logical processors) 10th gen i7 local machine, as no parallel processing is required.

Key assumptions of the theoretical model

1. Vaccination coverage at least increases over time (vaccination rollout is either pre-completed, or ongoing throughout the study)
2. A share of the population is vaccination-willing (*vax-willing*; conversely vaccination-unwilling or *vax-unwilling*) but may not necessarily receive the vaccine
3. A share of the population is test-willing (conversely, test-unwilling) but may not necessarily test for the disease
4. The vax-unwilling's probability of testing, suppose test-willing, may differ from that of the vax-willing
5. Vaccine effectiveness differs between individuals, but is distributed along a particular mean/mode
6. Suppose not yet infected, then the risk of infection follows a memoryless process
7. All individuals can be infected at most once, hence no re-infections
8. The baseline infection risk of the vax-unwilling may differ from that of the vax-willing

Hence, there are four groups — (1) Vax-willing and test-willing, (2) Vax-willing but test-unwilling, (3) Vax-unwilling but test-willing, (4) Vax-unwilling and test-unwilling

There are 10 parameters in the model (888125 valid combinations)

Parameter	Meaning	Values for Simulation
N	Simulated population size	1000
T	Number of time periods used in the simulation; this is equivalent to the follow-up duration	20, 30, 40, 50, 60
Θ_v	Fraction of population that is willing to be vaccinated (vax-willing); fraction Θ_v takes values $\theta_v = 1$ and fraction $1 - \Theta_v$ takes values $\theta_v = 0$	0.3, 0.4, 0.5, 0.6, 0.7
p_v	Probability of being vaccinated if not already vaccinated but vax-willing in period t; $p_v = 1$ implies static coverage, such that all vax-willing start off vaccinated, and $p_v \to 0$ the opposite	0.15, 0.3, 0.4, 0.5, 0.6, 0.85, 1
Θ_τ	Fraction of population that is willing to be tested (test-willing); fraction Θ_τ takes values $\theta_\tau = 1$ and fraction $1 - \Theta_\tau$ takes values $\theta_\tau = 0$	0.5, 0.65, 0.75, 0.85, 1
p_τ	Probability of being tested if test-willing in period t	0.5, 0.65, 0.75, 0.85, 1
k_τ	Ratio of probability of being tested of the vax-unwilling to the vax-willing ($\frac{p_\tau^{\theta_v=0}}{p_\tau^{\theta_v=1}}$); $k_\tau < 1$ indicates that the vax-unwilling are less likely to be tested than the vax-willing	0.5, 0.75, 0.9, 1, 1.1, 1.25, 1.5
α_b	Baseline attack rate for the vax-willing, which is the daily probability of contracting the disease if vax-willing but unvaccinated	0.025
k_α	Ratio of the baseline infection rate of the vax-unwilling relative to that of the vax-willing ($\frac{\alpha_b^{\theta_v=0}}{\alpha_b}$); $k_\alpha < 1$ indicates that the baseline infection rate of the vax-unwilling is lower than that of the vax-willing	0.5, 0.75, 0.9, 1, 1.1, 1.25, 1.5
μ_{VE}	Mean of the VE distribution, $VE_i \sim Beta(\alpha = 9, \beta = \frac{\alpha}{\mu_{VE}} - \alpha)$, $0 < \mu_{VE} < 1$ for individual i	0.5, 0.6, 0.7, 0.8, 0.9

Bias is decreasing in p_v, and increases as k_τ and k_α diverge from 1

$$Bias\{\widehat{VE}\} = \widehat{VE} - VE = \sum_{t=0}^{T} w_t \Big(\frac{\Theta_v(1-\Theta_v)\phi_v\phi'_{uv}(\Omega' - \Omega)}{(\Theta_v\phi_{uv} + (1-\Theta_v)\phi'_{uv})\,(\Theta_v\phi_{uv}\Omega + (1-\Theta_v)\phi'_{uv}\Omega')} \Big) \tag{1}$$

$$\textbf{where } VE = \sum_{t=0}^{T} w_t VE_t = \sum_{t=0}^{T} w_t \Big(1 - \frac{\Theta_v\phi_v}{\Theta_v\phi_{uv} + (1-\Theta_v)\phi'_{uv}}\Big) \tag{2}$$

$$\textbf{and } \widehat{VE} = \sum_{t=0}^{T} w_t \widehat{VE}_t = \sum_{t=0}^{T} w_t \Big(1 - \frac{\Theta_v\phi_v\Theta_\tau\Omega}{\Theta_v\phi_{uv}\Theta_\tau\Omega + (1-\Theta_v)\phi'_{uv}\Theta_\tau\Omega'}\Big) \tag{3}$$

$$\textbf{and } \phi_v = \sum_{m}^{t-1}(1-p_v)^{t-1-m}p_v \sum_{n}^{t-1}(1-\alpha_{uv})^n(1-\alpha_v)^{t-1-n}\alpha_v \tag{4}$$

$$\textbf{and } \phi_{uv} = \sum_{m}^{t}(1-p_v)^m \sum_{n}^{t-1}(1-\alpha_{uv})^n\alpha_{uv} \tag{5}$$

$$\textbf{and } \phi'_{uv} = \sum_{m}^{t-1}(1-k_\alpha\alpha_{uv})^m k_\alpha\alpha_{uv} \tag{6}$$

$$\textbf{and } \Omega = j(p_\tau) = \sum_{m}^{t} p_\tau \sum_{n}^{t-1}(1-p_\tau)^n p_\tau \sum_{q}^{t-1}(1-p_\tau)p_\tau^q \sum_{r}^{t-1}(1-p_\tau)^{t-1-r}p_\tau^r \tag{7}$$

$$\textbf{and } \Omega' = j(k_\tau p_\tau) = \sum_{m}^{t} k_\tau p_\tau \sum_{n}^{t-1}(1-k_\tau p_\tau)^n k_\tau p_\tau \sum_{q}^{t-1}(1-k_\tau p_\tau)k_\tau p_\tau^q \sum_{r}^{t-1}(1-k_\tau p_\tau)^{t-1-r}(k_\tau p_\tau)^r \tag{8}$$

Interpreting the theoretical bias equation

- Bias is a function of the behavioural asymmetries between the vax-willing and vax-unwilling (Θ_τ, p_τ, and k_τ), and the speed of vaccination rollout (p_v)
- Unambiguously, if the testing behaviour of the vax-willing and vax-unwilling are identical ($k_\tau = 1$, hence $\Omega = \Omega'$), then theoretical bias is nil
- The prevalence of test-willing individuals (Θ_τ) does not matter, but the prevalence of vax-willing (Θ_v) matters

Simulation focuses on commonly used cohorts and test-negative designs

Study Design	Estimation Approach	Notes
Cohort using aggregated data, with person-days as offset	Negative binomial regression (count analysis)	Commonly used where aggregate data is preferred or accessible in lieu of granular data, albeit often with a Poisson regression
Cohort using aggregated data, with daily population as offset	Negative binomial regression (count analysis)	Same as above but uses directly population, which may be time-varying, instead of person-days
Cohort using individual-level data, adjusting for immortal time bias	Cox Regression (survival analysis)	Often used in prospective cohorts where time-to-event is of interest, and where the quality of exposure-censoring time variables are robustly collected
Cohort using individual-level data	Logistic regression (binary outcomes)	Often used where time measures are unreliable, when time-to-event not of interest, as opposed to occurrence of outcomes, or when assumptions necessary for survival models, e.g., proportional hazards, fail
TND using individual-level data, keeping only the first positive, and, if never positive, the first negative test	Logistic regression (binary outcomes)	Often used when testing data from ambulatory care settings, healthcare facility screening, administrative registers, and ILI/SARI surveillance are deployed
TND using individual-level data, keeping only the first positive, but allowing for multiple negative tests	Logistic regression (binary outcomes)	Used when multiple follow-up points per unique individuals are available, such as through community infection surveys

Scenarios to illustrate bias of existing VE estimates in practice

Scenario	Approximant	p_v	Θ_τ	p_τ	k_τ	k_α	μ_{VE}
Scenario A (Slow rollout, comprehensive testing and tracing, vax-unwilling-biased baseline risks)	Wild Type	0.15	1	1	1	1.25	0.9
Scenario B (Slow rollout, partly comprehensive testing and tracing, vax-unwilling-biased baseline risks)	Alpha/ Beta/ Gamma	0.15	0.75	1	1	1.25	0.9
Scenario C (Gradual rollout, selective testing and tracing, symmetric baseline risks)	Delta	0.3	0.75	0.75	1	1	0.7
Scenario D (Fast rollout, highly selective testing and tracing, vax-willing-biased baseline risks)	Omicron BA1-2	0.5	0.75	0.5	0.75	0.75	0.5
Scenario E (Rapid rollout, minimal testing and tracing, vax-willing-biased baseline risks)	Omicron BA3-5	0.85	0.5	0.5	0.75	0.75	0.5
Scenario A & B (Severe Outcomes)	Wild Type/ Alpha/ Beta/ Gamma	0.15	1	1	1	1	0.9
Scenario C (Severe Outcomes)	Delta	0.3	1	1	1	1	0.9
Scenario D (Severe Outcomes)	Omicron BA1-2	0.5	1	1	1	1	0.9
Scenario E (Severe Outcomes)	Omicron BA3-5	0.85	1	1	1	1	0.9

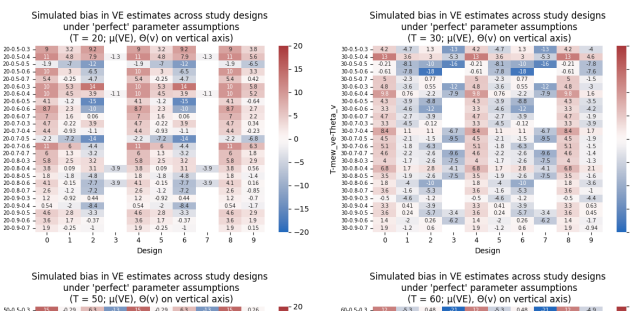
Estimate bias-parameter sensitivities using OLS linear regressions — (1) design-stratified, and (2) dummy fixed effects

Parameters are redefined to help interpret relative importance

- $k_\tau \rightarrow |k_\tau - 1|$; absolute deviation from 1 (deviation from **independent testing rates**)
- $p_v \rightarrow |p_v - 1|$; absolute shortfall from 1 (deviation from **static vaccine coverage**)
- $p_\tau \rightarrow |p_\tau - 1|$; absolute shortfall from 1 (deviation from **perfect testing if test-willing**)
- $\Theta_\tau \rightarrow |\Theta_\tau - 1|$; absolute shortfall from 1 (deviation from **full test-willing population**)
- $\Theta_v \rightarrow |\Theta_v - 1|$; absolute shortfall from 1 (deviation from **full vax-willing population**)
- $\mu_{VE} \rightarrow |\mu_{VE} - 1|$; absolute shortfall from 1 (deviation from **'perfect' vaccine**)

Invariable parameters (N, and α_b) are omitted* **'Ideal' conditions refers to when $p_v = 1$, $\Theta_\tau = 1$, $p_\tau = 1$, $k_\tau = 1$, and $k_\alpha = 1$

Under 'ideal' assumptions, survival analysis and count analysis cohorts, and TNDs with multiple negative tests perform best



	StudyDesign	Encoding
0	Cohort: True v	0
1	Cohort (immortal time correction): True v	1
2	Cohort (aggregate; PD): True v	2
3	Cohort (aggregate; N): True v	3
4	Cohort: Observed v	4
5	Cohort (immortal time correction): Observed v	5
6	Cohort (aggregate; PD): Observed v	6
7	Cohort (aggregate; N): Observed v	7
8	TND: First Pos, First Neg	8
9	TND: First Pos, Multiple Neg	9

Downward bias is sizeable in scenarios D and E, where deviations from 'ideal' are largest, and most representative of the recent Omicron waves

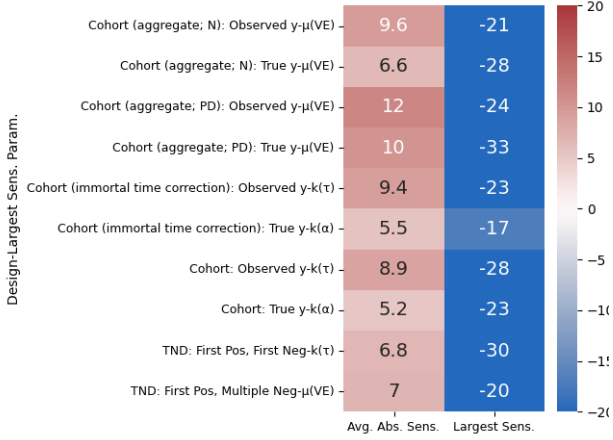
Wave	N	Theta_y	p_y	Theta_tau	p_tau	kitau	alpha_b	kialpha	asave_se	T	Cohort (aggregate; N): Observed y	Cohort (aggregate; N): True y	Cohort (aggregate; PD): Observed y	Cohort (aggregate; PD): True y	Cohort (nonnatal time correction): Observed y	Cohort (nonnatal time correction): True y	Cohort: Observed y	Cohort: True y	TND: First Pos, First Neg	TND: First Pos, Multiple Neg
Scenario A	1000	0.3	0.15	1	1	1	0.03	1.25	0.9	20			3.14	3.14	1.14	1.14	5.88	5.88	5.88	1.54
										30			1.85	1.85	0.55	0.55	5.7	5.7	5.7	0.84
										40			2.32	2.32	0.77	0.77	6.24	6.24	6.24	0.88
										50	-3.03	-3.03	1.93	1.93	2.62	2.62	7.35	7.35	7.35	2.62
										60	-9.12	-9.12	-3.89	-3.89	1.2	1.2	7.21	7.21	7.21	1
Scenario B	1000	0.3	0.15	0.75	1	1	0.03	1.25	0.9	20			-2.17		-0.99	0.67	1.16	4.77	5.79	5.51	0.95
										30	0.68	-1.54		1.54	1.97	0.27	4.92	5.27	6.14	1.81
										40	-2.01	-7.14	2.23	-10.55	3.3	0.94	5.84	6.07	7.11	3.49
										50	-7.93	-9.91	-10.41	-44.2	0.68	-0.1	5.97	6.07	6.82	0.41
										60	-4.05	-5.88	3.35	1.53	2.29	2.41	5.29	7.78	8.09	3.04
Scenario C	1000	0.5	0.3	0.75	0.75	1	0.03	1	0.7	20			-14.11	-14.11	-1.05	-1.58	6.04	9.8	7.86	-3.53
										30			-13.52	-14.41	-3.09	-3	2.46	9.8	5.37	-8.04
										40	-10.73		-55.61	-49.61	3.87	1.61	6.96	14.32	10.54	0.48
										50			-22.16	-21.89	0.63	0.12	2.12	13.79	6.86	-2.47
										60	-21.33	-19.58	-26.79	-23.32	-2.59	-1.89	-2.27	13.7	3.76	-6.39
Scenario D	1000	0.7	0.5	0.75	0.5	0.75	0.03	0.75	0.5	20	-11.14		-3.62	-29.37	-10.94	0.38	6.25	12.8	3.37	3.64
										30	-69.1		-87.47		-38.9	-37.12	6.68	-38.24	5.66	-84.74	-91.39
										40	-72.12	-19.69	-66.14		-53.14	-54.62	-8.22	-64.51	5.16	-61.41	-56.75
										50	-74.66	-31.88	-67.15		-60.13	-60.6	-22.72	-68.14	-6.92	-66.35	-61.26
										60	-89.88		-234.68		-91.21	-88.28	-27.93	-82.26	-12.15	-80.34	-80.06
Scenario E	1000	0.7	0.85	0.5	0.5	0.75	0.03	0.75	0.5	20			-54.49		-27.29	-33.51	-12.03	-20.64	-5.1	-34.23	-14.59
										30	-36.48	-19.83	-17.82		-27.56	-30.33	-13.59	-33.74	-5.43	-32.25	-14.04
										40	-64.52		-37.52		-15.54	-31.4	-12.18	-64.01	-2.57	-66.02	-34.32
										50		-30.96	91.73		-55.81	-50.43	-21.34	-60.53	-6.61	-61.92	-34.86
										60	-82.44	-20.26	-62.94		-81.35	-84.47	-16.78	-86.34	-3.37	-90.1	-52.05
Severe Outcomes (Scenario A & B)	1000	0.3	0.15	1	1	1	0.03	1	0.9	20	-0.42	-0.42	1.04	1.04	-1.25	-1.25	4.39	4.39	4.39	-0.78
										30	-2.18	-2.18	-0.18	-0.18	-2.06	-2.06	3.04	3.04	3.04	-1.67
										40	-6.39	-6.39	-5.34	-5.36	-3.7	-3.7	4.46	4.46	4.46	-1.34
										50	-4.9	-4.9	-0.03	-0.03	0.31	0.31	5.63	5.63	5.63	0.3
										60	-11.19	-11.19	-7.15	-7.15	-1.21	-1.21	5.41	5.41	5.41	-1.42
Severe Outcomes (Scenario C)	1000	0.5	0.3	1	1	1	0.03	1	0.9	20	-5.46	-5.46	-0.03	40.03	1.65	1.65	5.17	5.17	5.17	1.5
										30	-3.96	-3.96	-16.99	-16.99	0.28	0.28	4.7	4.7	4.7	0.15
										40	-5.74	-5.74	-11.46	-11.46	-0.54	-0.54	4.62	4.62	4.62	-0.02
										50			-20.41	-20.41	1.93	1.93	6.6	6.6	6.6	1.87
										60	-7.37	-7.37	-12.26	-12.26	0.96	0.96	6.42	6.42	6.42	0.74
Severe Outcomes (Scenario D)	1000	0.7	0.5	1	1	1	0.03	1	0.9	20			-8.02	-8.02	-0.03	-0.03	2.46	2.46	2.46	-1
										30			-8.65	-8.65	-2.06	-2.06	2.11	2.11	2.11	-2.29
										40	-4.39	-4.39	-14.32	-14.32	0.07	0.07	4.28	4.28	4.28	0.12
										50	-4.55	-4.55	-5.56	-5.56	0.95	0.95	5.4	5.4	5.4	0.9
										60	-10.8	-10.8	-8.18	-8.18	-1.04	-1.04	4.79	4.79	4.79	-2.21
Severe Outcomes (Scenario E)	1000	0.7	0.85	1	1	1	0.03	1	0.9	20			-3.6	-3.6	-0.49	-0.49	1.93	1.93	1.93	-0.18
										30			-1.08	-1.08	-1.38	-1.38	1.94	1.94	1.94	-1.18
										40	-3.25	-3.25	0.3	0.3	0.09	0.09	4.27	4.27	4.27	1.12
										50	-3.8	-3.8	0.38	0.38	1.45	1.45	5.4	5.4	5.4	1.58
										60	-9.8	-9.8	-1.16	-1.16	-1.24	-1.24	4.79	4.79	4.79	-1.3

Summary of findings on **drivers of bias**

- Across all study designs, the deviation in the mean of the VE distribution (μ_{VE}) from 'perfect' protection (100%) exerts a downward bias; this reflects the asymmetry in the beta distribution
- In binary outcomes and survival analysis cohorts, k_τ, and p_τ are most important. Of secondary importance are k_α, p_v, Θ_v, and Θ_τ
- In count analysis cohorts, k_τ, p_v, p_τ, and Θ_v are most important
- In TNDs with only one test per individual, only k_τ stood out, followed secondarily by p_τ, and k_α
  - As this configuration aims to remove the influence of timing of vaccination on the risk of disease exposure, p_v has a smaller association with bias; same goes forΘ_τ
- In the TND with multiple negative tests per individual, k_α is important, followed secondarily by k_τ, Θ_v, and p_v.
- **Both specifications of TNDs are more robust to 'less-than-ideal' conditions than cohorts, making TNDs attractive** when vaccine rollout is dynamic, and when testing is imperfect, such as during resource-constrained periods.
- *Immense gains if steps are taken to mitigate selective testing in respective infectious disease surveillance systems, such as through randomness in community sampling.*

Decomposition of simulated bias using dummy FE and stratified OLS regression shows vast heterogeneity in the parameter-bias nexus

Average and Largest Sensitivity Between Parameters Ξ and Absolute Bias (Parameter With Largest Sensitivity on Vertical Axis)



Discussion

1. Study design choices have **non-trivial effects on VE estimation bias**
2. Findings warrant a **re-benchmarking of methodology**
   - Recommendations for VE research need to reflect COVID-19-specific nuances
   - Use heat maps of bias-parameter sensitivities to 'adjust' VE estimates
   - Future research ought to layout the sources of bias discussed to help form judgment on the extent of estimation bias $\rightarrow$ **additional reporting checklist**
3. Helps design cost-effective surveillance systems by quantifying part of the **estimation bias vs. implementation cost trade-off**
   - Conceptually, the policymaker 'chooses' a system that (1) mitigates sources of bias, and (2) geared towards the least biased VE study design, given resource constraint
   - Resource-constrained countries may opt to concentrate their surveillance systems on major outbreak zones to use TNDs, rather than more resource-intensive cohorts
4. **Disease progression, or severity of symptoms, that is correlated with testing propensity** can be discussed within the model
   - Low testing rate amongst the test-willing ($p_\tau \ll 1$) + lower testing rate amongst the vax-willing ($k_\tau \gg 1$) $\rightarrow$ symptoms tend to be milder amongst the vaccinated, hence less likely to test

We recommend future VE research to report the following, in addition to the STROBE checklist

No.	Item	Purpose
1	Pace of vaccination (daily time series or average) **during** the study period, including trend breaks, e.g., surge in doses given	To gauge p_v
2	Vaccine coverage **by the end of** the study period	To gauge Θ_v (and potentially k_v)
3	The pace of vaccination rollout **after** the study period	To gauge Θ_v (and potentially k_v)
4	Testing rate in the study population, and the overall population	To gauge Θ_τ and p_τ
5	Testing rate by vaccination status or relevant comparator groups	To gauge k_τ
6	Contact tracing regime (if relevant for disease studied) in place during the study period, e.g., comprehensive forward and backward, symptomatic forward only, no contact tracing	To gauge Θ_τ, and p_τ
7	Testing regime in place during the study period, including scope, eligibility, barriers of accessibility (direct and indirect), and selection criteria related to vaccination, e.g., universal (including asymptomatic) testing with private cost, symptomatic testing with private cost, symptomatic testing with price or queue discrimination on vaccination status	To gauge Θ_τ, p_τ, and k_τ
8	Presence of NPIs to influence vaccine take-up via restrictions on disease transmission-relevant activities	To gauge k_α
9	**[For TNDs]** Baseline characteristics of study participants stratified by vaccination status (usually reported by test-positive, and test-negative)	To gauge k_α, and k_τ
10	**[Where available]** VE estimates from meta-analyses, and systematic reviews specific to the disease, pathogenic variant, and clinical outcome of interest	To form priors on μ_{VE} beyond the study

Limitations

1. Tests with sub-100% sensitivity and specificity is not explicitly modelled
   - Reasonable sensitivities and specificities (in the region of 80% to 90%) $\rightarrow$ minimal bias for both cohorts and TNDs (*Jackson and Rothman (2015)*)
2. Re-infections are not explicitly modelled
   - Past infections compound with vaccines $\rightarrow$ undetected infections amongst vaccinated $>$ unvaccinated $\rightarrow$ overestimate VE
3. Observed confounding is not explicitly modelled
   - Requires strong and non-generalisable assumptions on the structure between potential confounders, vaccination choice, and disease

Additional Findings

(Part 1/2) Unless indicated otherwise, range of colours $= [-20, 20]$

- A hypothetical 20 ppt over-estimation represents a substantial gap in estimated, and 'true' protection offered.
- A 20 ppt over-estimation puts this the estimated COVID-19 vaccine effectiveness against death from a meta-analysis of 8 vaccine platforms spanning 5 VOCs (Higdon et al, 2022) from the neighbourhood of 60% - 99% to 40% - 79%.
- The equivalent for severe disease goes from 55% - 99% to 35% - 79%, and that of SARS-CoV-2 infection to well below 50%.

(Part 2/2) Unless indicated otherwise, range of colours = $[-20, 20]$

- For sufficiently low baseline attack rates ($< 1\%$), as is for severe COVID-19, 20 ppt can represent sizeable changes in NNV even when the estimated VE is high

Figure: Contour Map of NNV against VE, Conditional on α_b

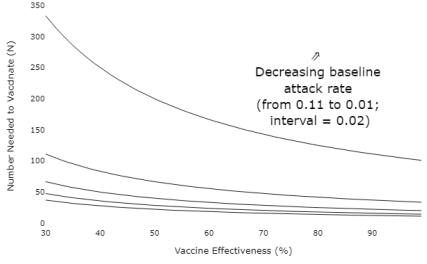
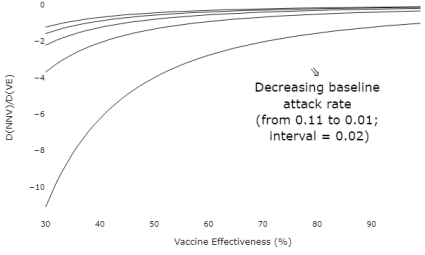
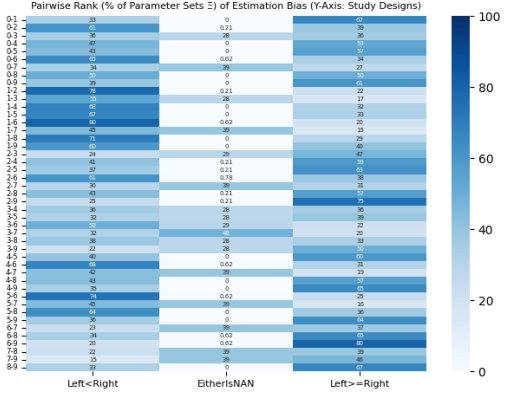


Figure: Contour Map of first derivative of NNV against VE, Conditional on α_b



Rank reversals are common; these rankings need to consider real-world frequency of the corresponding parameter settings



	StudyDesign	Encoding
0	Cohort: True y	0
1	Cohort (immortal time correction): True y	1
2	Cohort (aggregate; PD): True y	2
3	Cohort (aggregate; N): True y	3
4	Cohort: Observed y	4
5	Cohort (immortal time correction): Observed y	5
6	Cohort (aggregate; PD): Observed y	6
7	Cohort (aggregate; N): Observed y	7
8	TND: First Pos, First Neg	8
9	TND: First Pos, Multiple Neg	9

Rescaled (1) Association Between Parameters Ξ and Absolute Bias

(**Scenario D**: $p_v = 0.5$; $\Theta_\tau = 0.75$; $p_\tau = 0.5$; $k_\tau = 0.75$; $k_\alpha = 0.75$)

Rescaled (2) Association Between Parameters Ξ and Absolute Bias

(**Scenario D + Mitigation of Selection in Testing**: $p_V = 0.5$; $\Theta_\tau = 0.75$; $p_\tau = 0.75$; $k_\tau = 0.95$; $k_\alpha = 0.75$)

FE estimates of drivers of bias (average within-design sensitivities)