How Important Are Study Designs? A Simulation Assessment of Vaccine Effectiveness Estimation Bias with Time-Varying Vaccine Coverage, and Heterogeneous Testing and Baseline Attack Rates

Jing Lian Suah[a†*]   Naor Bar-Zeev[b]   Maria Deloria Knoll[b]

[a] Research and Modelling Unit, Central Bank of Malaysia, Malaysia

[b] Bloomberg School of Public Health, Johns Hopkins University, United States

[†] This research was conducted while on secondment at the Institute for Clinical Research, National Institutes of Health, Ministry of Health, Malaysia

[*] Corresponding author; email address: suahjinglian@bnm.gov.my

August 26, 2022

Abstract

Objective. We studied how commonly used vaccine effectiveness (VE) study designs (variations of cohorts, and test-negative designs) perform under epidemiological nuances more prominent in the COVID-19 era, specifically time-varying vaccine coverage, and heterogeneous testing behaviour and baseline attack rates with selection on willingness to vaccinate.

Methodology. We simulated data from a multi-parameter conceptual model of the epidemiological environment using 888125 parameter sets. Four configurations of cohorts, and two test-negative designs, were conducted on the simulated data, from which estimation bias is computed. Finally, stratified and fixed effects linear regressions were estimated to quantify the sensitivity of estimation bias to model parameters.

Findings. Irrespective of study designs, dynamic vaccine coverage, and heterogeneous testing behaviour and baseline attack rates are important determinants of bias. Study design choices have non-trivial effects on VE estimation bias even if these factors are absent. The importance of these sources of bias differ across study designs.

Conclusion. A re-benchmarking of methodology, especially for studying COVID-19 VE, and implementation of vaccine-preventable disease surveillance systems that minimise these sources of bias, are warranted.

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.

- There is intrinsic estimation bias across all study designs, with the direction and magnitude contingent on specific conditions.

- 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 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.

1 Introduction

When first-generation COVID-19 vaccines were rolled out via mass vaccination in early 2021, the World Health Organisation (WHO) issued an interim guidance on evaluating vaccine effectiveness (VE), based primarily on past non-COVID-19 outbreaks (WHO 2021). Reflecting this, the WHO made an explicit preference for test-negative (TND) or test-negative case-control (TNCC) designs in minimising confounding from health-seeking behaviour, while acknowledging the merits of retrospective and prospective cohorts. Other methods, including the screening method, and regression discontinuity design (RDD), were discussed (Orenstein et al. 1985, WHO 2021). This recommendation, and the makeup of the COVID-19 VE literature, provide the basis to evaluate variations of cohorts and TNDs.

The pre-COVID-19 literature discussed several issues pertinent to the estimation bias of VE. Firstly, under reasonable sub-100% testing sensitivities and specificities, both cohorts, and case-controls and TNDs showed small biases under simulations (Orenstein et al. 2007), but correction for misclassification bias is possible (Patel et al. 2020, Endo et al. 2020). This holds for COVID-19, given the regulations requiring high accuracy rates in professionally admin-

istered Reverse Transcription Polyamerase Chain Reaction (RT-PCR) and Rapid Antigen (RTK-Ag) tests. Secondly, TNDs are argued to be more robust to selection on health-seeking behaviour than cohorts (Jackson & Nelson 2013). Thirdly, TNDs are unbiased under a range of assumptions, but temporary non-specific immunity post-infection of similar non-target diseases, asymmetric severity between the vaccinated and unvaccinated, and, imbalances in the rates of non-target diseases in the same symptom cluster (e.g., influenza-like illnesses that are influenza and non-influenza) between the vaccinated and unvaccinated result in non-trivial biases (Foppa et al. 2013, Jackson & Nelson 2013, Jackson et al. 2018). Fourthly, study designs, when constructing counterfactuals, ought to consider different models of vaccine failure (e.g., 'leaky' vaccines, waning VE), confounders (e.g. time-varying exposure risk, and disease predisposition), and outcome definitions (Crowcroft & Klein 2018). The COVID-19 VE literature correctly uses specific designs to assess waning relative to static VE, and regression analysis to address confounders. Attention was also given to outcome definitions, such as differences between post-COVID-19-positive deaths and clinically audited deaths due to COVID-19, and between clinically confirmed SARS-CoV-2 infections and positive tests.

Notwithstanding the aforementioned concerns, the COVID-19 vaccine rollout in 2021-22 deviated in substantial respects from pre-pandemic vaccination distribution. Firstly, pre-COVID-19 vaccine take-up tend to be stable, but not during the COVID-19 pandemic (Ritchie et al. 2020). A COVID-19-era simulation noted the importance of temporal effects arising from differential risks at equivalent post-vaccination time points (Lewnard et al. 2021). However, even with constant background risk, those vaccinated earlier have more opportunities of disease exposure post-vaccination than those vaccinated later. Secondly, high but stagnant coverage without long-run supply constraints indicate selection between the vaccinated and unvaccinated, potentially on health-seeking, testing, and risk behaviour (Ritchie et al. 2020). Thirdly, testing regimens varied, ranging from selective to comprehensive (Ritchie et al. 2020). These may present new sources of bias for COVID-19 VE studies amid vaccine rollouts.

This study asks "how do commonly used VE study designs (variations of cohorts, and TNDs) perform amid (1) time-varying vaccine coverage, (2) latent selection of and risk behaviours on willingness to vaccinate and (3) 'leaky' vaccines (such that $VE < 100\%$ against infection)?" We attempt to answer in three steps. Firstly, a theoretical discussion of estimation bias. Secondly, a simulation of the theorised environment, and VE estimates under variations of cohorts and TNDs. Thirdly, an analysis of the drivers of estimation bias.

2 Methods

2.1 Conceptual Framework and Simulation

Data are simulated from a conceptual framework with 10 adjustable parameters. There are four behavioural blocks, detailed in Appendix A — (1) vaccination take-up, (2) testing behaviour, (3) distribution of vaccine effectiveness, and (4) disease outcomes. The following parameters choices Ξ were used, amounting to 888125 unique combinations. We considered range for generalisability, and sparsity for computational ease. 'Ideal' conditions are when vaccine coverage is static, and behavioural frictions in testing and baseline risk do not exist $(p_v = 1,\ \Theta_\tau = 1,\ p_\tau = 1,\ k_\tau = 1,\ k_\alpha = 1)$.

Parameter	Meaning	Values for Simulation
N	Simulated population size	1000
T	Number of chronological days 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

Hence, there are four groups.

1. Unwilling to be vaccinated (vax-unwilling) and tested (test-unwilling): $\theta_v = 0$ and $\theta_\tau = 0$

2. Unwilling to be vaccinated (vax-unwilling), but willing to be tested (test-willing) : $\theta_v = 0$ and $\theta_\tau = 1$

3. Willing to be vaccinated (vax-willing), but unwilling to be tested (test-unwilling): $\theta_v = 1$ and $\theta_\tau = 0$

4. Willing to be vaccinated (vax-willing) and tested (test-unwilling): $\theta_v = 1$ and $\theta_\tau = 1$)

Theoretical bias is as follows (equations 1 to 8). w_t refers to the daily weights implicit to the choices of study designs, which has no closed form. Full derivation, including special cases of static vaccine coverage ($T = 1$), is in Appendix B.

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

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

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

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

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

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

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

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

Ceteris paribus, theoretical bias decreases in p_v (upward bias), and increases as k_τ (downward) and k_α (upward) diverge from 1. Direction is inferred from signs in equation 1. Theoretical bias primarily depends the behavioural asymmetries between those who are vax-willing and vax-unwilling, and the speed of vaccination rollout. While the incidence of vax-willing individuals enters the equation, incidence of test-willing individuals does not, and only testing behaviour does. An unbiased design, given parameter set Ξ, optimises for w_t, such that theoretical bias is zero.

$$w_t | \Xi \equiv \{\sum_{t=0}^{T} w_t \hat{VE}_t = \mathbb{E}(VE)\} \tag{9}$$

2.2 Study Designs

We analyse two families of study designs — (i) cohorts and (ii) test-negative designs (TNDs) — for their ubiquity and merits in the literature when handling two pertinent issues.

First, the asymmetry in testing propensity arising from $k_\tau \neq 1$. TNDs seek to address this by conditioning the study population on having being tested, hence health-seeking behaviour (Crowcroft & Klein 2018). Cohorts can control for baseline testing behaviour, e.g., number of tests taken pre-observation.

Second, time-varying vaccine coverage when $T > 1$ and/or $p_v < 1$. Unlike pre-COVID-19 studies, COVID-19 vaccines are evaluated amid mass vaccination. From the conceptual framework, bias arises from that amongst the vaccinated, those vaccinated earlier have more draws of disease exposure till the end of time T than those vaccinated later. Ceteris paribus, breakthrough infection probability is higher for early-takers. In cohorts, researchers adjust for time-at-risk (e.g, using person-days or dynamic population as offsets in count models, using time-to-event via survival models, and controlling for timing of disease exposure or vaccination). In TNDs, the researcher may restrict to only the first positive, or first negative (if never positive) test, rather than multiple negatives per individual. These are analogous to solving for daily weights w_t in the conceptual model such that the probability of testing positive between time of vaccination t_v and the end of study T is independent of time.

$$Pr(\sum_{t=t_v}^{T} \hat{y}_{i,t} = 1 | t_v = k, v = 1) > Pr(\sum_{t=t_v}^{T} \hat{y}_{i,t} = 1 | t_v = j, v = 1) \; for \; k < j \tag{10}$$

$$w_t \equiv \{Pr(\sum_{t=t_v}^{T} \hat{y}_{i,t} = 1 | t_v = k, v = 1) \perp t_v, t \; \forall \; 1 \leq k \leq T\} \tag{11}$$

Table 2 summarises the study designs. For cohorts, we further consider when true outcomes are observable. TNDs require testing.

Table 2: Permutations of Study Designs

Study Design	Approach Used to Estimate VE	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 (Bar-On et al. 2021, Goldberg et al. 2021, Suah, Husin, Tok, Tng, Thevananthan, Low, Appannan, Zin, Zin, Yahaya et al. 2022)
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 (Jara et al. 2021, n.d.)
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 (Suah, Husin, Tok, Tng, Thevananthan, Low, Appannan, Zin, Zin, Yahaya et al. 2022, Suah et al. 2021)
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 (Olson et al. 2022)
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 (Bernal et al. 2021)

For all parameter sets Ξ, the simulated data are used to estimate the VE, β is the relative risk (negative binomial regression), hazard ratio (Cox regression), or odds ratio (logistic regression). Bias is the difference between estimated $\hat{VE}$ and mean of the VE distribution μ_{VE}. Absolute bias ($|\hat{VE} - \mu_{VE}|$) is also used.

$$\widehat{VE} = 100\% * (1 - exp(\widehat{\beta_v})) \tag{12}$$

$$Bias(\widehat{VE}) = \widehat{VE} - \mu_{VE} \tag{13}$$

2.3   Scenario-Specific Parameter Assumptions

To illustrate estimation bias in practice, we isolate parameter sets that reflect various historical SARS-CoV-2 waves (A: Wild Type; B: Alpha / Beta / Gamma; C: Delta; D: Omicron, BA1-2; E: Omicron, BA3-5), such as the supply and demand of vaccines (p_v), contact tracing and testing practices (Θ_τ, p_τ, k_τ), and behavioural context of those unwilling to be vaccinated (k_α). Detailed justifications are in the supplementary appendix. As surveillance of severe outcomes (hospitalisation, ICU admission, and death) is often stricter due to more specific outcome definitions, and clinical audits, these corresponding scenarios assume frictionless testing ($\Theta_\tau = 1$, $p_\tau = 1$, $k_\tau = 1$) and symmetric baseline risk behaviour ($k_\alpha = 1$).

Table 3: Scenario-Specific Parameter Assumptions

	Θ_v	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)	0.3	0.15	1	1	1	1.25	0.9
Scenario B (Slow rollout, partly comprehensive testing and tracing, vax-unwilling-biased baseline risks)	0.3	0.15	0.75	1	1	1.25	0.9
Scenario C (Gradual rollout, selective testing and tracing, symmetric baseline risks)	0.5	0.3	0.75	0.75	1	1	0.7
Scenario D (Fast rollout, highly selective testing and tracing, vax-willing-biased baseline risks)	0.7	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)	0.7	0.85	0.5	0.5	0.75	0.75	0.5
Scenario A & B (Severe Outcomes)	0.3	0.15	1	1	1	1	0.9
Scenario C (Severe Outcomes)	0.5	0.3	1	1	1	1	0.9
Scenario D (Severe Outcomes)	0.7	0.5	1	1	1	1	0.9
Scenario E (Severe Outcomes)	0.7	0.85	1	1	1	1	0.9

2.4   Drivers of Bias

To assess the relative importance of the specific parameters on simulated bias, we estimated the sensitivity of estimation bias to parameter values using two linear regressions with the ordinary least squares (OLS) estimator. Firstly, by study designs. Secondly, the average 'within-design' sensitivities using dummy fixed effects (FE). Technical aspects are in Appendix C. As the FE estimator returns a weighted average measure of association, divergent signs between designs are masked.

To ease interpretation, variable parameters are transformed as follows. To limit influence from outliers, estimation biases outside $[-60, 60]$ are excluded, leaving behind 886413 parameter sets. Invariable parameters (N, and α_b) are omitted.

- $k_\tau \rightarrow |k_\tau - 1|$; 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)

- $\Theta_\tau \to |\Theta_\tau - 1|$; absolute shortfall from 1 (deviation from full test-willing population)

- $\Theta_v \to |\Theta_v - 1|$; absolute shortfall from 1 (deviation from full vax-willing population)

- $\mu_{VE} \to |\mu_{VE} - 1|$; absolute shortfall from 1 (deviation from 'perfect' vaccine)

3   Findings

The simulation took 415516 seconds (115.42 hours) in Python 3.10 on a AWS EC2 c6i.32xLarge Amazon Linux instance with 128 virtual central processing units, one of the most powerful public-accessible compute-optimised virtual machines when the analysis was conducted.

The colour range in all heat maps of estimation biases is constrained to $[-20, 20]$. A 20 percentage point (ppt) over-estimation puts the estimated COVID-19 VE against death from a meta-analysis of 8 platforms spanning 5 variants of concern (Higdon et al. 2022) from the neighbourhood of 60% - 99% to 40% - 79%. The equivalent for severe disease from 55% - 99% to 35% - 79%, and that of SARS-CoV-2 infection to below 50%. This hypothetical 20 ppt over-estimation represents a substantial gap in estimated, and 'true' protection offered. A contour map of number needed to vaccinate (NNV) and VE, given any input of the baseline attack rate α_0, is also available for inspection in the supplementary appendix. For sufficiently low baseline attack rates ($< 1\%$), as is likely the case for severe COVID-19, a 20 ppt change in VE represents sizeable change in NNV, even if VE is high. E.g., changing VE from 90% to 70% when the baseline attack rate is 1% raises the NNV from 100 to 150, a 50% increase in NNV. The colour scheme differentiating over- (red) and under-estimation (blue) reflects that over-estimation ($Bias(\widehat{VE}) > 0$) may be more damaging, as policymakers may be overconfident in unwinding public health restrictions.

3.1   Simulation

The simulated performance of each design differs across parameter sets Ξ (full table of 888125 parameter combinations is here in supplementary materials; a snippet is in the appendix). Under 'ideal' conditions ($p_v = 1$, $\Theta_\tau = 1$, $p_\tau = 1$, $k_\tau = 1$, $k_\alpha = 1$; figure 1), some designs are equivalent. Under perfect testing, cohorts where tests are required to observe disease outcomes and where tests are not required are identical. If testing is comprehensive, the TND with single test contribution is identical to the binary outcomes cohort. In the 'ideal' case, the TND that allows for multiple negative tests is the least biased, followed closely by the survival analysis cohort. The count analysis cohort with population as the offset underestimates, while the binary outcomes cohort, and TND with single test contribution overestimates. As follow-up lengthens (T) under 'ideal' conditions, estimation biases of the

binary outcome cohort, single test TND, and count analysis cohort with population as the offset increase.

Under gradual rollout ($p_v \neq 1$, all else ideal), the opposite holds. Relaxing instead perfect testing ($\Theta_\tau < 1$ and/or p_τ and/or $k_\tau \neq 1$, all else ideal), simulated biases change in a mixed manner. Lowering only the population share of test-willing Θ_τ reverses some findings. Next, some relative performances are reversed if pre-completion of rollout, and perfect testing are relaxed ($p_v < 1$, $\Theta_\tau < 1$, $p_\tau < 1$, $k_\tau \neq 1$). Asymmetry in baseline attack rates between the vax-willing and vax-unwilling ($k_\alpha \neq 1$) tend to increase the degree of bias. Higher baseline attack rates for the vax-unwilling relative to the vax-willing tend to overstate VE, while the opposite tend to understate. Without long-run supply constraints, everyone who is willing to be vaccinated will be vaccinated. If long-run vaccine coverage is stable but not 100%, the vaccinated and unvaccinated may be fundamentally different in behaviour. Hence, studies comparing the two groups under high vaccine coverage may be biased.

Figure 1: Simulated estimation bias under 'ideal' parameter assumptions

(a)

(b)

(c)

(d)

(e)

(f)
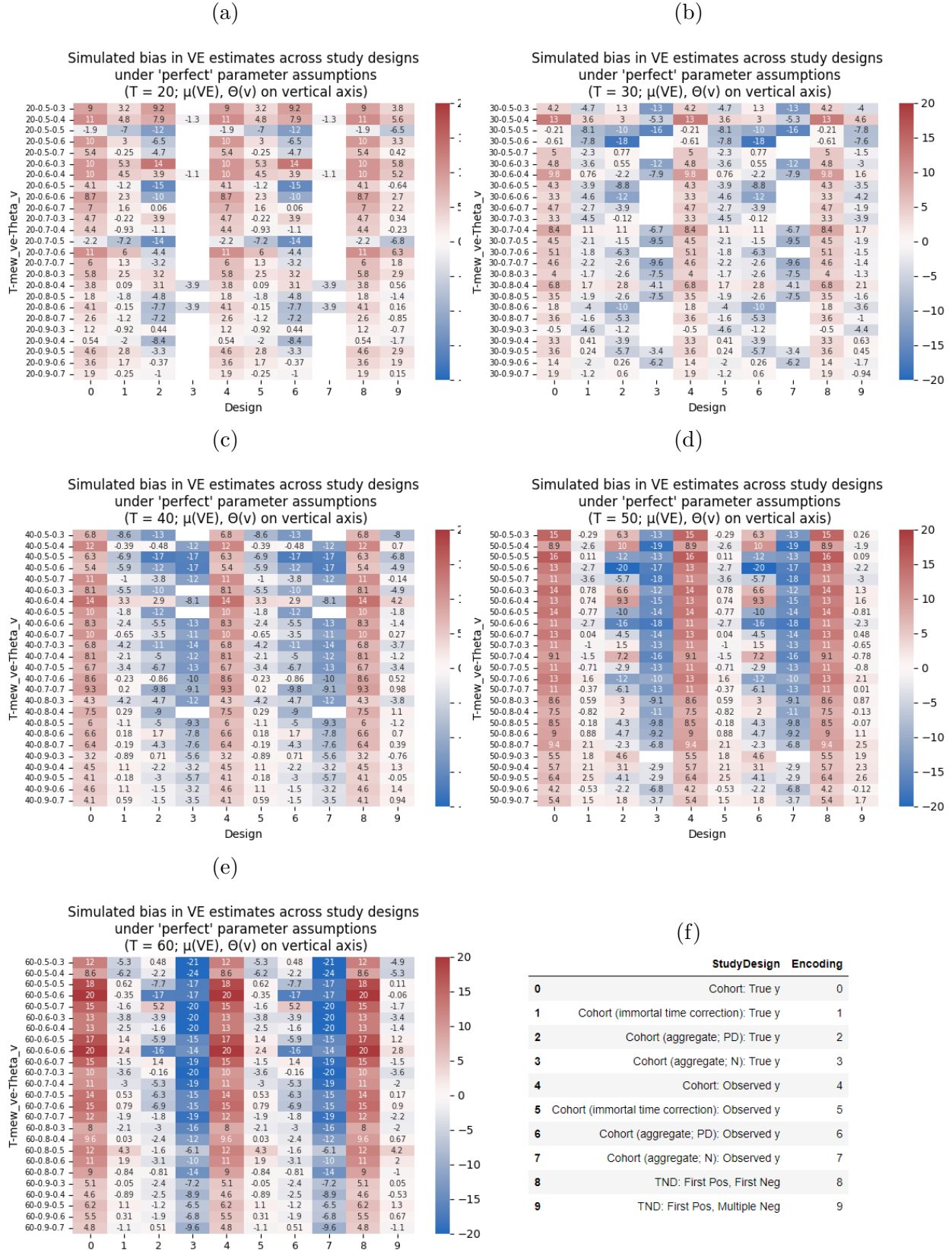
The estimation biases vary under all nine scenarios-specific parameter sets (figure 2). Count analysis cohorts were generally biased downwards. In scenario A where parameters were least deviated from 'ideal', biases were low. In scenarios B and C, survival analysis cohorts, and TNDs with multiple negative tests performed consistently better. In scenario C, binary outcomes cohort were comparable to the former designs, but with an upward bias. In scenarios D and E, the TND with multiple negative tests was marginally least biased, and downward bias is sizeable across all designs. As frictionless testing, and symmetric baseline risk ($\Theta_\tau = 1$, $p_\tau = 1$, $k_\tau = 1$, $k_\alpha = 1$) are assumed in the analogues of severe outcomes, the only source of bias is the speed of vaccination rollout (p_v). Hence, estimation bias is minimal except for the count analysis cohorts. For completeness, the supplementary appendix formally analyses the stability of relative ranks between pairwise designs.

Figure 2: Simulated estimation bias under scenario-specific parameter sets

Wave	N	Theta_v	p_v	Theta_tau	p_tau	ktau	alpha_b	kalpha	mew_ve	T	Cohort (aggregate; N): Observed y	Cohort (aggregate; N): True y	Cohort (aggregate; PD): Observed y	Cohort (aggregate; PD): True y	Cohort (immortal time correction): Observed y	Cohort (immortal 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	3.7	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	1.54	1.97	0.37	4.92	5.37	6.14	1.51
										40	-2.01	-7.14	2.21	-10.55	3.5	0.94	5.84	6.07	7.11	3.49
										50	-7.93	-9.91	-40.48	-44.3	0.68	-0.1	3.97	6.67	6.82	0.41
										60	-4.05	-5.88	3.35	1.33	3.29	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			-34.51	-34.59	-1.05	-1.38	6.04	9.8	7.86	-3.53
										30			-13.52	-14.48	-3.89	-3	2.48	9.37	5.57	-4.76
										40	-10.73		-55.61	-49.61	3.87	1.61	6.98	14.32	10.54	0.45
										50			-22.16	-23.89	0.63	0.12	2.12	13.79	6.86	-3.67
										60	-21.23	-19.58	-26.78	-22.32	-2.59	-1.89	-2.27	13.7	3.76	-6.59
Scenario D	1000	0.7	0.5	0.75	0.5	0.75	0.03	0.75	0.5	20	-13.04		-3.62	-29.07	-10.06	0.58	-4.25	12.8	-3.27	5.04
										30	-49.1		-87.47	-38.9	-37.12	-6.68	-36.34	5.66	-34.74	-18.29
										40	-72.12	-19.69	-66.14	-53.14	-54.62	-9.25	-64.51	5.16	-61.41	-36.75
										50	-74.05	-31.88	-177.15	-80.13	-60.6	-22.72	-69.34	-6.92	-66.35	-41.28
										60	-89.88		-234.68	-94.21	-68.26	-27.83	-82.26	-12.15	-50.54	-40.06
Scenario E	1000	0.7	0.85	0.5	0.5	0.75	0.03	0.75	0.5	20			-54.49	-37.29	-33.81	-12.03	-35.64	-5.1	-34.23	-14.96
										30	-36.48	-19.83	-17.82	-37.56	-30.35	-13.59	-33.76	-5.49	-22.25	-14.04
										40	-64.52		-37.53	-13.54	-5.4	-12.18	-64.81	-2.57	-66.02	-34.32
										50		-28.96	-91.75	-55.81	-30.63	-21.24	-80.75	-10.55	-61.92	-34.66
										60	-52.44	-28.26	-62.94	-41.05	-44.47	-16.94	-56.34	-5.37	-50.1	-32.05
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	-1.28	-1.28	4.39	4.39	4.39	-0.76
										30	-2.18	-2.18	-0.18	-0.18	-2.06	-2.06	3.84	3.84	3.84	-1.67
										40	-6.39	-6.39	-5.36	-5.36	-1.7	-1.7	4.46	4.46	4.46	-1.51
										50	-4.9	-4.9	-0.03	-0.03	0.51	0.51	5.65	5.65	5.65	0.5
										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	-40.03	-40.03	1.65	1.65	5.17	5.17	5.17	1.5
										30	-3.96	-3.96	-16.89	-16.89	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	-0.81
										50			-20.41	-20.41	1.93	1.93	6.6	6.6	6.6	1.87
										60	-7.97	-7.97	-12.26	-12.26	0.96	0.96	6.42	6.42	6.42	0.76
Severe Outcomes (Scenario D)	1000	0.7	0.5	1	1	1	0.03	1	0.9	20			-8.02	-8.02	-0.83	-0.83	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	4.26	4.26	4.26	0.12
										50	-4.55	-4.55	-5.56	-5.56	0.98	0.98	5.4	5.4	5.4	0.9
										60	-10.8	-10.8	-8.18	-8.18	-1.84	-1.84	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	-1.16
										40	-3.25	-3.25	0.3	0.3	0.89	0.89	4.37	4.37	4.37	1.12
										50	-3.8	-3.8	0.36	0.36	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

3.2 Drivers of Bias

Figure 3 summarises the sensitivity of estimation bias to parameters by study design. Figure 4 shows the average sensitivity, and the parameter with the largest sensitivity. The appendix contains a version with sensitivities rescaled to scenario D ($p_v = 0.5$; $\Theta_\tau = 0.75$; $p_\tau = 0.5$; $k_\tau = 0.75$; $k_\alpha = 0.75$), and another version that reflects scenario D with steps taken to mitigate testing frictions ($p_v = 0.5$; $\Theta_\tau = 0.75$; $p_\tau = 0.75$; $k_\tau = 0.95$; $k_\alpha = 0.75$) to gauge potential gains for policymakers if surveillance mitigate selective testing, and the FE model estimates as a summative exercise. This 'rescaling' assists researchers who often lack choice on methodology, a direct result of pre-existing surveillance and data structures, gauge the estimation bias they are working with.

Across all designs except the binary outcomes cohort without testing requisites, deviation in the mean of the VE distribution (μ_{VE}) from 'perfect' protection (100%) exerts a downward
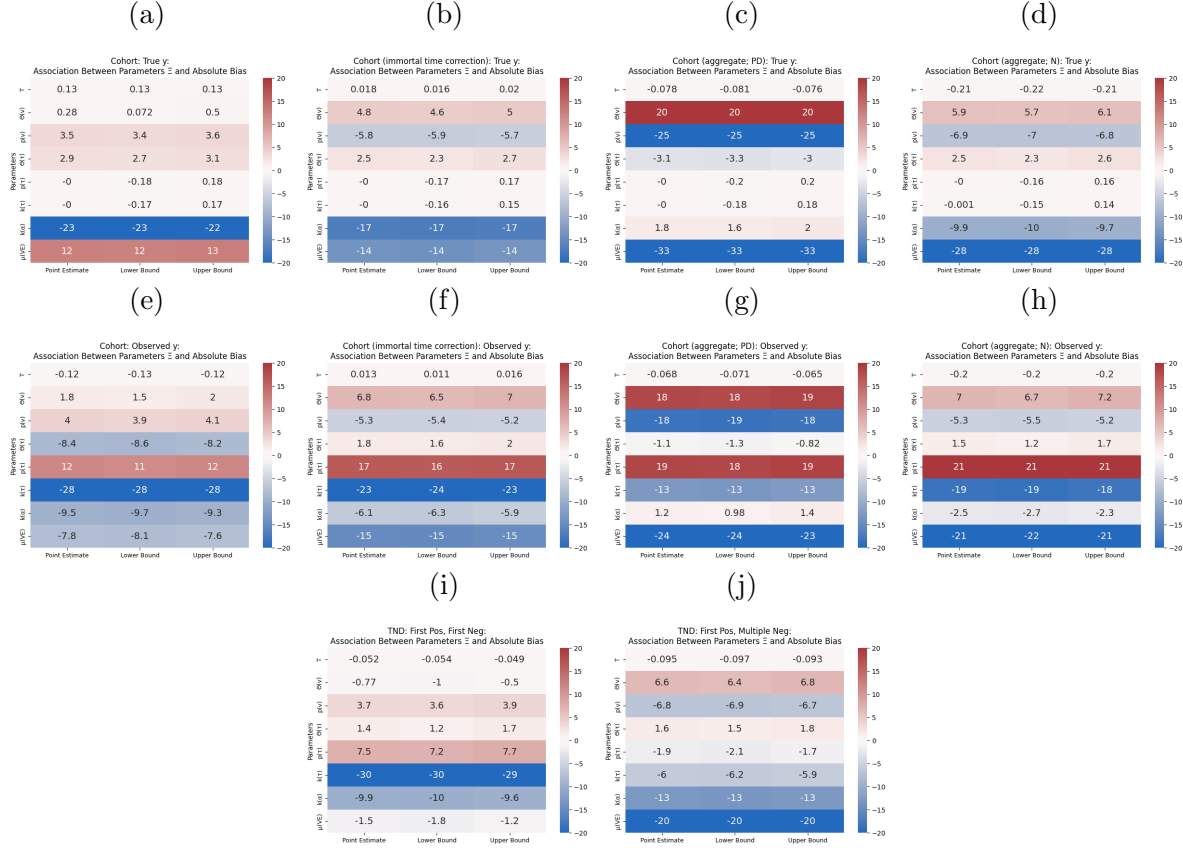
bias. This reflects asymmetry in the beta distribution, as all estimators (MLE on negative binomial, logistic, and Cox regressions) require log-normal errors, which could be unrealistic in reality.

When true outcomes are observable (figures 3a to 3d), in binary outcomes and survival analysis cohorts, the asymmetry in baseline attack rates between the vax-willing and vax-unwilling (k_α) is important. In the latter, deviation from static vaccine coverage (vaccination rate, p_v) is salient. In count analysis, cohorts with person-days as offset, the deviation from static vaccine coverage (vaccination rate, p_v; downward bias) is most important, but secondary to the asymmetry in baseline attack rates between the vax-willing and vax-unwilling (k_α) in the case with population as offset.

When testing is prerequisite to observe outcomes (figures 3e to 3h), in binary outcomes and survival analysis cohorts, the asymmetry in testing propensity between the vax-willing and vax-unwilling (k_τ), and the deviation from perfect testing amongst the test-willing population (p_τ) are most important. Of secondary importance are the asymmetry in baseline attack rates between the vax-willing and vax-unwilling (k_α), the deviation from static vaccine coverage (vaccination rate, p_v), full test-willing population (Θ_τ), and full vax-willing population (Θ_v). In count analysis cohorts, the asymmetry in testing propensity between the vax-willing and vax-unwilling (k_τ), the deviation from static vaccine coverage (vaccination rate, p_v), perfect testing amongst the test-willing population (p_τ) and full vax-willing population (Θ_v) are most important.

In TNDs with only one test per individual (figure 3i), only the asymmetry in testing propensity between the vax-willing and vax-unwilling (k_τ) stood out, followed by the deviation from perfect testing amongst the test-willing population (p_τ), and the asymmetry in baseline attack rates between the vax-willing and vax-unwilling (k_α). As this configuration aims to remove the influence of timing of vaccination on the risk of disease exposure, the deviation from static vaccine coverage (vaccination rate, p_v) has a small association with bias, similarly for deviation from full test-willing population (Θ_τ). In contrast, asymmetry in baseline attack rates between the vax-willing and vax-unwilling (k_α) is important in the TND with multiple negative tests per individual (figure 3j), followed by the asymmetry in testing propensity between the vax-willing and vax-unwilling (k_τ), and the deviation from full vax-willing population (Θ_v), and static vaccine coverage (vaccination rate, p_v).

Figure 3: Association Between Parameters Ξ and Absolute Bias By Study Design

Finally, the average absolute sensitivity between parameters and absolute bias indicate that, amongst study designs where testing is prerequisite to observing outcomes, both specifications of TNDs are most robust to 'less-than-ideal' conditions. These properties make the TND attractive under dynamic vaccine rollout, and imperfect testing. The largest sensitivity measure is roughly equal across all designs.

Figure 4: Average Absolute and Largest Sensitivity Between Parameters Ξ and Absolute Bias (Parameter With Largest Sensitivity on Vertical Axis)
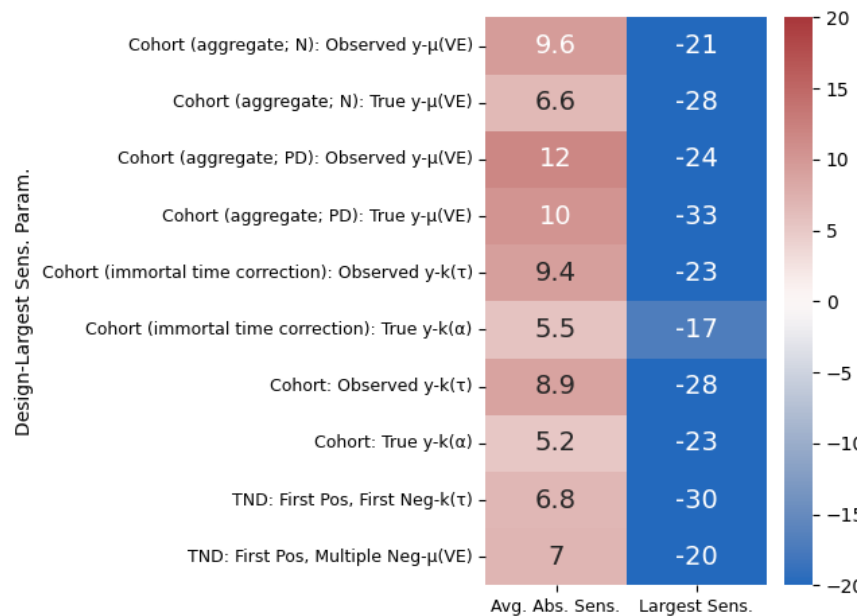
4 Discussion

Where vaccine rollouts are gradual, testing is heterogeneous, and baseline risk associated with vaccination attitudes, study design choices have non-trivial effects on VE estimation bias. This extends a pre-COVID-19 review that noted study design may be important in uncovering specific models of vaccine failures and underlying mechanisms (Crowcroft & Klein 2018). Our research is unique in that it formalises, without loss of generality, the interaction between study design and the epidemiological environment in affecting estimation bias.

We modelled directly the epidemiological nuances of COVID-19 vaccines governing the pace of rollout, preference heterogeneity in testing, risk behaviour, and vaccination, and uncertainty in unobservable individual-level VE. Open access and official data globally indicate these properties (Ritchie et al. 2020). Pre-COVID-19, vaccine coverage remained stable over years, such as rotavirus vaccines (Hungerford et al. 2018). Where coverage is stable and high, such as BCG, Pol3, DTP3, and IPV (Vanderslott et al. 2013), those unwilling to be vaccinated may be systematically different from those willing. Coverage of these vaccines are similar to COVID-19 vaccines. Balanced take-up, such as PCV3, YFV, and Rotavirus vaccines (Vanderslott et al. 2013), may indicate supply constraints, such that selection in baseline risk and testing on vaccination preference may be less prominent.

Our study simulated data from a multi-parameter conceptual model of the epidemiological

environment, reflective of when COVID-19 vaccines were rolled out. Hence, we can pin down the effects of (1) parameters, and (2) study designs on estimation bias. Suppose real-world data is used, both will be mixed with multidirectional selection bias. Moreover, a straightforward decomposition of bias allows researchers to understand the design-specific parameter-bias nexus. The analysis is generalisable to the increasing number of studies on effectiveness relative to vaccinated subgroups, e.g., primary vaccinated (Suah, Tng, Tok et al. 2022, Bar-On et al. 2021, Ranzani et al. 2022), and single booster vaccinated (Regev-Yochay et al. 2022). The 'baseline' group for behavioural frictions is simply the intended reference, rather than the unvaccinated. Going forward, the same principle applies to studies comparing next-generation vaccines, or variant-updated vaccines against existing first-generation vaccines.

Our work has three implications.

Firstly, they warrant a re-benchmarking of methodology. Recommendations for VE research, such as by the WHO (WHO 2021), need to reflect these considerations, as recent novel diseases are materially different from earlier ones during which most designs were evaluated. Our analysis of the drivers of estimation bias confirm that (1) the parameter-bias nexus differs across designs, and (2) some designs are more robust to sources of bias. In practice, researchers can use the heat maps of bias-parameter sensitivities to 'adjust' VE estimates, by first ascertaining the design used, and second calibrate parameters in the conceptual model, e.g., a study during the Omicron wave should have large deviations from 'ideal' values in the testing parameters, but small deviations in vaccination rate. Confidence intervals of the VE estimates can be expanded this way. Alternatively, researchers may adjust the point estimate and the confidence interval concurrently uni-directionally. For readers of empirical research, VE estimates may be adjusted ex post to assess where the 'true' VE may lie, e.g., VE estimates during the Omicron wave may require substantial upward adjustments, or wider confidence bands, than those from the Delta wave. Hence, to help form judgment on the extent of estimation bias, we recommend future research to report the following, in addition to the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) checklist.

Table 4: Checklist for VE Research to Gauge Estimation Bias

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

Secondly, this study informs policymakers on designing cost-effective vaccine-preventable infectious disease surveillance systems by quantifying part of the estimation bias versus implementation cost trade-off, crucial for resource-constrained countries, or when outbreaks require temporary, cheap, surveillance. While count and survival analyses cohorts perform best under 'ideal' circumstances, TNDs are more resilience against sources of bias. Countries

may gear surveillance in major outbreaks to use TNDs, rather than cohorts, which require more comprehensive data infrastructures. Conceptually, the policymaker 'chooses' a surveillance system that (1) explicitly mitigates sources of bias, and (2) enables the least biased study design, subject to a resource constraint. Both (1) and (2) are dynamically related, as some designs perform better under 'ideal' circumstances, but not otherwise. This motivates a formal analysis on cost-effective design optimisation. Moreover, given the importance of testing and risk behaviour heterogeneity in VE estimation bias, surveillance systems may include community random sampling, e.g., the United Kingdom's COVID-19 Infection Survey (ONS 2022).

Thirdly, disease progression, or severity of symptoms, that is correlated with testing propensity can be discussed. In principle, this entails another layer of outcome post-infection with a unique relationship with vaccination status and testing preferences. However, this requires an excessively complex theoretical model with minimal gains in insights. Within the current model, this is best captured via testing rate (p_τ) and asymmetry (k_τ). Combinations of low testing rate amongst the test-willing ($p_\tau \ll 1$), and low testing propensity amongst the vax-willing than the vax-unwilling ($k_\tau \gg 1$) represent instances where symptoms are milder for the vaccinated, who are then less likely to test unless sufficiently symptomatic.

There are limitations in this study. Firstly, tests with sub-100% sensitivity and specificity is not modelled. However, a pre-COVID-19 simulation found that reasonable sensitivities and specificities (80% to 90%) yielded minimal bias (Jackson & Rothman 2015). Secondly, reinfections are not explicitly modelled. Suppose past infections provide some protective effect that compound with that of vaccines, then greater prevalence of undetected infections amongst the vaccinated than the unvaccinated may overestimate VE. An example is if breakthrough infections are extremely mild but close to each other, such that viral loads are sufficiently low to not trigger antigen tests, then estimated VE is high when true VE is low. This warrants more involved modelling in future research. Thirdly, observed confounders are not modelled explicitly. Doing so requires strong and non-generalisable assumptions on the structure between confounders, vaccination choice, and disease dynamics. Fourthly, waning VE (or building of VE) are also not modelled explicitly, which could exponentiate the complexity of the simulation. Further research may consider various forms of non-linearity in waning, interaction with behavioural frictions, and probabilistic waning.

5   Conclusion

Using simulated data from a multi-parameter conceptual model of the epidemiological environment, we offer three findings. Firstly, both dynamic vaccine coverage, and heterogeneous testing behaviour and baseline attack rates that are selected on willingness to vaccinate,

19

are important determinants of VE estimation bias. Secondly, study designs have non-trivial effects on estimation bias, even without these frictions. Thirdly, these factors have different relative and absolute importance in driving estimation bias. Moving forward, a re-benchmarking of methodology, and surveillance systems of vaccine-preventable disease that minimise these sources of bias, are warranted.

Ethical consideration

As this research uses only simulated data, and does not involve any human participants, institutional review board (IRB) approval is not required.

Declaration of interests

Data and replication statement

This research uses only simulated data, and does not involve any human participants. All scripts required for replication are available at **github.com/suahjl/vemethod-simulation-frictions**.

Acknowledgement

References

Bar-On, Y. M., Goldberg, Y., Mandel, M., Bodenheimer, O., Freedman, L., Kalkstein, N., Mizrahi, B., Alroy-Preis, S., Ash, N., Milo, R. et al. (2021), 'Protection of BNT162b2 vaccine booster against COVID-19 in Israel', *New england journal of medicine* **385**(15), 1393–1400.

Bernal, J. L., Andrews, N., Gower, C., Gallagher, E., Simmons, R., Thelwall, S., Stowe, J., Tessier, E., Groves, N., Dabrera, G. et al. (2021), 'Effectiveness of COVID-19 vaccines against the B. 1.617. 2 (Delta) variant', *New England Journal of Medicine* .

Crowcroft, N. S. & Klein, N. P. (2018), 'A framework for research on vaccine effectiveness', *Vaccine* **36**(48), 7286–7293.

Endo, A., Funk, S. & Kucharski, A. J. (2020), 'Bias correction methods for test-negative designs in the presence of misclassification', *Epidemiology & Infection* **148**.

Foppa, I. M., Haber, M., Ferdinands, J. M. & Shay, D. K. (2013), 'The case test-negative design for studies of the effectiveness of influenza vaccine', *Vaccine* **31**(30), 3104–3109.

Goldberg, Y., Mandel, M., Bar-On, Y. M., Bodenheimer, O., Freedman, L., Haas, E. J., Milo, R., Alroy-Preis, S., Ash, N. & Huppert, A. (2021), 'Waning immunity after the BNT162b2 vaccine in Israel', *New England Journal of Medicine* **385**(24), e85.

Higdon, M. M., Wahl, B., Jones, C. B., Rosen, J. G., Truelove, S. A., Baidya, A., Nande, A. A., ShamaeiZadeh, P. A., Walter, K. K., Feikin, D. R., Patel, M. K., Deloria Knoll, M. & Hill, A. L. (2022), 'A Systematic Review of Coronavirus Disease 2019 Vaccine Efficacy and Effectiveness Against Severe Acute Respiratory Syndrome Coronavirus 2 Infection and Disease', *Open Forum Infectious Diseases* **9**(6). ofac138.
URL: *https://doi.org/10.1093/ofid/ofac138*

Hungerford, D., Vivancos, R., Read, J. M., Iturriza-Gómara, M., French, N. & Cunliffe, N. A. (2018), 'Rotavirus vaccine impact and socioeconomic deprivation: an interrupted time-series analysis of gastrointestinal disease outcomes across primary and secondary care in the UK', *BMC medicine* **16**(1), 1–14.

Jackson, M. L. & Nelson, J. C. (2013), 'The test-negative design for estimating influenza vaccine effectiveness', *Vaccine* **31**(17), 2165–2168.

Jackson, M. L., Phillips, C. H., Benoit, J., Kiniry, E., Madziwa, L., Nelson, J. C. & Jackson, L. A. (2018), 'The impact of selection bias on vaccine effectiveness estimates from test-negative studies', *Vaccine* **36**(5), 751–757.

Jackson, M. L. & Rothman, K. J. (2015), 'Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness', *Vaccine* **33**(11), 1313–1316.

Jara, A., Undurraga, E. A., González, C., Paredes, F., Fontecilla, T., Jara, G., Pizarro, A., Acevedo, J., Leo, K., Leon, F. et al. (2021), 'Effectiveness of an inactivated SARS-CoV-2 vaccine in Chile', *New England Journal of Medicine* **385**(10), 875–884.

Jara, A., Undurraga, E. A., Zubizarreta, J. R., Gonzalez, C., Pizarro, A., Acevedo, J., Leo, K., Paredes, F., Bralic, T., Vergara, V. et al. (n.d.), 'Effectiveness of homologous and heterologous booster shots for an inactivated SARS-CoV-2 vaccine: A large-scale observational study'.

Lewnard, J. A., Patel, M. M., Jewell, N. P., Verani, J. R., Kobayashi, M., Tenforde, M. W., Dean, N. E., Cowling, B. J. & Lopman, B. A. (2021), 'Theoretical framework for retrospective studies of the effectiveness of SARS-CoV-2 vaccines', *Epidemiology (Cambridge, Mass.)* **32**(4), 508.

Olson, S. M., Newhams, M. M., Halasa, N. B., Price, A. M., Boom, J. A., Sahni, L. C., Pannaraj, P. S., Irby, K., Walker, T. C., Schwartz, S. P. et al. (2022), 'Effectiveness of BNT162b2 vaccine against critical COVID-19 in adolescents', *New England Journal of Medicine* .

ONS (2022), 'Coronavirus (COVID-19) infection survey: Methods and further information'.
    URL: *https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseas*

Orenstein, E. W., De Serres, G., Haber, M. J., Shay, D. K., Bridges, C. B., Gargiullo, P. & Orenstein, W. A. (2007), 'Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness', *International journal of epidemiology* **36**(3), 623–631.

Orenstein, W. A., Bernier, R. H., Dondero, T. J., Hinman, A. R., Marks, J. S., Bart, K. J. & Sirotkin, B. (1985), 'Field evaluation of vaccine efficacy', *Bulletin of the World Health Organization* **63**(6), 1055.

Patel, M. M., Jackson, M. L. & Ferdinands, J. (2020), 'Postlicensure evaluation of COVID-19 vaccines', *JAMA* **324**(19), 1939–1940.

Ranzani, O. T., Hitchings, M. D., de Melo, R. L., de França, G. V., Cássia de Fátima, R. F., Lind, M. L., Torres, M. S. S., Tsuha, D. H., David, L. C., Said, R. F. et al. (2022), 'Effectiveness of an inactivated COVID-19 vaccine with homologous and heterologous boosters against the Omicron (B. 1.1. 529) variant', *medRxiv* .

Regev-Yochay, G., Gonen, T., Gilboa, M., Mandelboim, M., Indenbaum, V., Amit, S., Meltzer, L., Asraf, K., Cohen, C., Fluss, R. et al. (2022), 'Efficacy of a fourth dose of COVID-19 mRNA vaccine against Omicron', *New England Journal of Medicine* **386**(14), 1377–1380.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D. & Roser, M. (2020), 'Coronavirus pandemic (COVID-19)', *Our world in data* .

Suah, J. L., Husin, M., Tok, P. S. K., Tng, B. H., Thevananthan, T., Low, E. V., Appannan, M. R., Zin, F. M., Zin, S. M., Yahaya, H. et al. (2022), 'Waning COVID-19 vaccine effectiveness for BNT162b2 and CoronaVac in Malaysia: an observational study', *International Journal of Infectious Diseases* **119**, 69–76.

Suah, J. L., Tng, B. H., Tok, P. S. K. et al. (2022), 'Real-world effectiveness of homologous and heterologous BNT162b2, CoronaVac, and AZD1222 booster vaccination against Delta and Omicron SARS-CoV-2 infection', *Emerging Microbes & Infections* **11**(1), 1343–1345. PMID: 35499301.
URL: *https://doi.org/10.1080/22221751.2022.2072773*

Suah, J. L., Tok, P. S. K., Ong, S. M., Husin, M., Tng, B. H., Sivasampu, S., Thevananthan, T., Appannan, M. R., Muhamad Zin, F., Mohd Zin, S. et al. (2021), 'PICK-ing Malaysia's epidemic apart: Effectiveness of a diverse COVID-19 vaccine portfolio', *Vaccines* **9**(12), 1381.

Vanderslott, S., Dadonaite, B. & Roser, M. (2013), 'Vaccination', *Our World in Data* .

WHO (2021), Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021, Technical report, World Health Organization.

A   Conceptual Framework (Theory)

The key assumptions of the conceptual framework are as follows.

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. **The vax-willing, and test-willing** — will get tested, and/or vaccinated at some probability to be determined

2. **The vax-willing, but test-unwilling** — will get vaccinated at some probability, but will never get tested

3. **The vax-unwilling, but test-wiling** — will get tested at some probability, but will never get vaccinated

4. **The vax-unwilling, and test-unwilling** — will never get vaccinated, nor get tested

These assumptions allow us to zoom into the aspects highlighted in the study question — (i) time-varying vaccine coverage, latent selection of (ii) testing, and (iii) risk behaviour, specifically on attitude towards vaccination.

A.1 Vaccination Take-Up

Over T time periods (e.g., days), $n_{v,t}$ are vaccinated on day t. By the end of time, $t = T$, $N_{v,t}$ are vaccinated, and $N_{uv,t}$ are unvaccinated.

$$N_t = \sum_{t=1}^{T}(n_{v,t}) + (N_t - \sum_{t=1}^{T}(n_{v,t})) = N_{v,t} + N_{uv,t} \tag{A.1}$$

Whether an individual i subscribes to the vaccination drive depends on a the vax-willingness parameter $(\theta_{v,i})$. The population share of individuals who are vax-willing is Θ_v. Vax-unwilling individuals never get vaccinated. This parameter is determined ex ante (before $t = 1$). Conditional on being vax-willing $(\theta_{v,i} = 1)$ and not already vaccinated, the probability of being vaccinated at period t is p_v. How close p_v is to 1 therefore determines the pace of the rollout, with $p_v = 1$ corresponding to a static coverage throughout $t = 1$ till $t = T$.

$$0 < \Theta_v < 1 \tag{A.2}$$
$$Pr(v_{i,t} = 1 | v_{i,t-1} = 0, \theta_{v,i} = 1) = p_v, \ 0 < p_v < 1 \tag{A.3}$$
$$Pr(v_{i,t} = 1 | \theta_{v,i} = 0) = 0 \tag{A.4}$$

A.2 Testing Behaviour

A fraction Θ_τ of the population is willing to undergo testing (test-willing). Test-willing individuals are indicated by their test-willingness parameter $\theta_\tau = 1$ (test-unwilling take $\theta_\tau = 0$). Individual i, suppose test-willing, will undergo testing at period t at probability p_τ. This process is, similarly, memoryless. Test-unwilling individuals never get tested.

$$0 < \Theta_\tau < 1 \tag{A.5}$$
$$Pr(\tau_{i,t} = 1 | \theta_{\tau,i} = 1) = p_\tau, \ 0 < p_\tau < 1 \tag{A.6}$$
$$Pr(\tau_{i,t} = 1 | \theta_{\tau,i} = 0) = 0 \tag{A.7}$$

However, the probability of testing amongst vax-unwilling but test-willing individuals $(\theta_{v,i} = 0$ and $\theta_\tau = 1)$ is a fraction $k_\tau \leq 1$ that of the vax-willing and test-willing $(\theta_{v,i} = 1$ and $\theta_\tau = 1)$

$$Pr(\tau_{i,t} = 1 | \theta_{v,i} = 0, \theta_{\tau,i} = 1) = k_\tau Pr(\tau_{i,t} = 1 | \theta_{v,i} = 1, \theta_{\tau,i} = 1) = k_\tau p_\tau, \ k_\tau \leq 1 = 0 \quad \text{(A.8)}$$

A.3 Distribution of Vaccine Effectiveness

The motivation of modelling VE as a random variable stems from immunogenic studies suggest that both humoral (S-antibody, and B-cell) and cellular (T-cell) responses against SARS-CoV-2 differ between individuals. While no formal correlates of protection against SARS-CoV-2 have been established, it is a reasonable assumption that VE is dynamic. From an estimation perspective, statistical methods then recover the central measure (often mean) of VE in a study population.

The VE for individual i is a random variable that follows a Beta distribution with a mean of μ_{VE}, bounded by 0 and 1, and with the α parameter preset to 9. A Beta distribution has two desirable properties — (1) bounded distribution, as, in theory, non-inferior vaccines have intrinsic $VE \in [0\%, 100\%]$, and (2) smooth but variable slopes (a strong contender is the Triangular distribution, which is smooth, bounded, but has invariant slopes). The choice of $\alpha = 9$ is arbitrary, and chosen for its effect on the concentration around the mean and mode; other researchers can choose higher α values for a more 'concentrated' distribution, and lower alpha values for the opposite. $\mu_{VE} = 0.5$ (VE of 50%) corresponds to a special case where the mode of the VE distribution equates its mean. Individual i, if vaccinated, will experience an attack rate $\alpha_{v,i}$ that is lower than if unvaccinated α_{uv} by the drawn VE. This constructs individual-specific counterfactuals. For simplicity, the attack rate for the unvaccinated α_{uv} is a scalar, i.e., constant for everyone that is unvaccinated.

$$VE_i \sim Beta(\alpha = 9, \beta = \frac{\alpha}{\mu_{VE}} - \alpha), \ 0 < \mu_{VE} < 1 \quad \text{(A.9)}$$

$$\alpha_{v,i} = VE_i \alpha_{uv}, 0 < \alpha_{uv} < 1 \quad \text{(A.10)}$$

A.4 Disease Outcome

The infection risk of individual i (α_i) is a function of vax-willingness $\theta_{v,i}$, vaccination status $v_{i,t}$, and the latent individual-specific vaccine effectiveness VE_i, which activates if a vaccine is administered. α_{base} refers to the baseline risk for vax-willing individuals ($\theta_{v,i} = 1$), and is scaled by k_α for the vax-unwilling ($\theta_{v,i} = 0$). If vaccinated ($v_i = 1$), then the infection risk for individual i is further scaled by the effectiveness of the vaccine $1 - VE_i$. k_α is therefore the ratio of baseline infection risk of the vax-unwilling to that of the vax-willing. Moreover, we

assume no re-infections. If not yet infected, the risk of infection (attack rate) for individual i is memoryless.

$$\alpha_i = \alpha(\theta_{v,i}, v_{i,t}, VE_i) = \alpha_{base} \cdot k_\alpha(\theta_{v,i}) \cdot (1 - VE_i(v_i)), \ k_\alpha > 0 \tag{A.11}$$
$$Pr(y_{i,t} = 1|y_{i,t-1} = 1) = 0 \tag{A.12}$$

Hence, there are three permutations of infection risks, depending on vax-willingness ($\theta_{v,i}$) vaccination status ($v_{i,t}$). (1) $\alpha_{vw',uv} = k_\alpha \alpha_{uv}$ for the vax-unwilling, and unvaccinated, (2) $\alpha_{vw,uv} = \alpha_{uv}$ for the vax-willing, but unvaccinated, and (3) $\alpha_{vw,v} = \alpha_v$ for the vax-willing, and vaccinated. The vax-unwilling, and vaccinated combination is paradoxical, hence does not exist, $\alpha_{vw',v} = k_\alpha \alpha_v = 0$.

$$Pr(y_{i,t} = 1|y_{i,t-1} = 0, \theta_{v,i} = 0, v_{i,t} = 0) = \alpha_{vw',uv} = k_\alpha \alpha_{uv} \tag{A.13}$$
$$Pr(y_{i,t} = 1|y_{i,t-1} = 0, \theta_{v,i} = 1, v_{i,t} = 0) = \alpha_{vw,uv} = \alpha_{uv} \tag{A.14}$$
$$Pr(y_{i,t} = 1|y_{i,t-1} = 0, \theta_{v,i} = 1, v_{i,t} = 1) = \alpha_{vw,v} = \alpha_v \tag{A.15}$$
$$Pr(y_{i,t} = 1|y_{i,t-1} = 0, \theta_{v,i} = 0, v_{i,t} = 1) = \alpha_{vw',v} = k_\alpha \alpha_v = 0 \tag{A.16}$$

To observe infection $\hat{y}$ at period t, individual i needs to be tested at the period of infection $\tau_{i,t} = 1$. We can generalise to a longer detection period, but this is unnecessary convolution for the aims of the study. Tests are assumed to be perfectly sensitive and specific in this framework.

Hence, if an individual i is infected at period t but does not take a test, the observed state will not indicate infection $\hat{y}_{i,t} = 0$. If an individual is indeed not infected at period t, taking a test will indicate a lack of infection $\hat{y}_{i,t} = 0$. In contrast, taking a test at period t when infected at period t will observe an infection $\hat{y}_{i,t} = 1$.

$$(\hat{y}_{i,t}|\tau_{i,t} = 1) = y_{i,t} \tag{A.17}$$
$$(\hat{y}_{i,t}|\tau_{i,t} = 0) = 0 \tag{A.18}$$

Therefore, absent of a universally test-willing population with a 100% testing rate, observed infection $\sum_t^T \sum_i^N \hat{y}_{i,t}$ is likely a fraction of true infection $\sum_t^T \sum_i^N y_{i,t}$.

$$\sum_{t}^{T}\sum_{i}^{N}\hat{y}_{i,t} = k_y \sum_{t}^{T}\sum_{i}^{N} y_{i,t} \ for \ \Theta_\tau < 1 \ and \ p_\tau < 1, \ where \ 0 < k_y < 1 \qquad \text{(A.19)}$$

B Theoretical Vaccine Effectiveness and Bias

VE is a function of the attack rate amongst the vaccinated relative the unvaccinated. Empirically, this is calculated from estimates of hazard ratios (survival analysis), odds ratios (binary outcomes), and relative rates (count analysis). Hence, bias is the difference between observed and true (calculated from true infection y) VE.

$$VE = 1 - \frac{Pr(y=1|v=1)}{Pr(y=1|v=0)} \qquad \text{(B.1)}$$

$$\widehat{VE} = 1 - \frac{Pr(y=1|v=1,\tau=1)}{Pr(y=1|v=0,\tau=1)} \qquad \text{(B.2)}$$

$$Bias\{\widehat{VE}\} = \widehat{VE} - VE \qquad \text{(B.3)}$$

We explore the theoretical bias of VE as the following assumptions are incrementally introduced — (1) differences in baseline attack rate between the vax-willing, and vax-unwilling, (2) differences in test-willingness, and selection in testing on vax-willingness, as well as (3) gradual rollout of vaccines.

B.1 Static: Selection in Risk on Vax-Willingness

Drawing from our conceptual model, the unvaccinated group is comprised of individuals who are willing to be vaccinated, but not vaccinated (e.g., due to supply constraints), as well as individuals who are unwilling to be vaccinated, whose baseline attack rates differ. Unbiased VE is recovered by comparing between the vaccinated and unvaccinated, conditional on being vax-willing. This is analogous to randomising the vaccinated and unvaccinated arms, such that there is no self-selection into either arms based on willingness to be vaccinated. Absent of this adjustment, the unvaccinated group is the average of the vax-willing and vax-unwilling, weighted by population share, whose baseline attack rates may differ due to fundamental differences in risk behaviour. Hence, a vanilla comparison of the vaccinated and unvaccinated arms in the presence of differential risk behaviour between the vax-willing and vax-unwilling results in biased VE estimates.

$$VE = 1 - \frac{Pr(y=1|v=1, \theta_v=1)}{Pr(y=1|v=0, \theta_v=1)} = 1 - \frac{\alpha_{vw,v}}{\alpha_{vw,uv}} = 1 - \frac{\alpha_v}{\alpha_{uv}} \tag{B.4}$$

$$\widehat{VE} = 1 - \frac{Pr(\hat{y}=1|v=1)}{Pr(\hat{y}=1|v=0)} = 1 - \frac{\alpha_v}{\Theta_v \alpha_{uv} + (1-\Theta_v)k_\alpha \alpha_{uv}} \tag{B.5}$$

$$Bias\{\widehat{VE}\} = \frac{\alpha_v}{\Theta_v \alpha_{uv} + (1-\Theta_v)k_\alpha \alpha_{uv}} - \frac{\alpha_v}{\alpha_{uv}} = \frac{\alpha_v}{\alpha_{uv}}\left(\frac{1}{\Theta_v + (1-\Theta_v)k_\alpha} - 1\right) \tag{B.6}$$

B.2 Static: Selection in Risk and Testing on Vax-Willingness

For simplicity, we begin with the special case where there is only one time period ($T=1$), but adding heterogeneity in testing rates based on vax-willingness. True VE is hence a function of the population share of vax-willing individuals, the probability of being vaccinated, and attack rates. The observed VE then further depends on the population share of test-willing, and the probability of being tested, as well as the relative testing propensity of the vax-unwilling to the vax-willing. Due to the asymmetry in the testing propensity and baseline attack rates between the vax-willing and vax-unwilling, the bias of observed VE is not 0. Bias is only zero if both testing propensity and baseline attack rates are independent of vax-willingness.

$$VE = 1 - \frac{\Theta_v p_v \alpha_v}{\Theta_v(1-p_v)\alpha_{uv} + (1-\Theta_v)k_\alpha \alpha_{uv}} = 1 - \frac{A}{B + k_\alpha C} \tag{B.7}$$

$$\widehat{VE} = 1 - \frac{\Theta_v p_v \alpha_v \Theta_\tau p_\tau}{\Theta_v(1-p_v)\alpha_{uv}\Theta_\tau p_\tau + (1-\Theta_v)k_\alpha \alpha_{uv}\Theta_\tau k_\tau p_\tau} = 1 - \frac{A\Theta_\tau p_\tau}{B\Theta_\tau p_\tau + Ck_\alpha \Theta_\tau k_\tau p_\tau} \tag{B.8}$$

$$Bias\{\widehat{VE}\} = \frac{\Theta_\tau p_\tau}{\Theta_\tau p_\tau}\left(\frac{A}{B + k_\tau k_\alpha C}\right) - \frac{A}{B + k_\alpha C} = \frac{A}{B + k_\tau k_\alpha C} - \frac{A}{B + k_\alpha C} \tag{B.9}$$

$$Bias\{\widehat{VE}\} = 0 \; if \; k_\tau = 1 \; and \; k_\alpha = 1 \tag{B.10}$$

B.3 Dynamic

In the dynamic case, there is more than one period ($T > 1$). We can think of overall VE as a weighted sum of period-specific VE_t. Weights depend on the t-specific share of the subgroups. The specific functional form is not necessary to analyse bias in observed overall $\widehat{VE}$.

Period-specific VE_t, which is expressed in terms of actual outcome y, is a function of vaccination rate (probability of receiving the vaccine, if vax-willing and not already vaccinated)

p_v, population share of vax-willing individuals Θ_v, the baseline attack rate α_{uv}, and the ratio of the baseline attack rate amongst the vax-unwilling relative to the vax-willing k_α.

$$VE_t = 1 - \frac{A_t}{B_t + C_t} \tag{B.11}$$

$$A_t = \Theta_v \underbrace{\sum_m^{t-1}(1 - p_v)^{t-1-m}p_v \sum_n^{t-1}(1 - \alpha_{uv})^n(1 - \alpha_v)^{t-1-n}\alpha_v}_{\phi_v = f(\alpha_v, \alpha_{uv}, p_v)} \tag{B.12}$$

$$B_t = \Theta_v \underbrace{\sum_m^t(1 - p_v)^m \sum_n^{t-1}(1 - \alpha_{uv})^n\alpha_{uv}}_{\phi_{uv} = g(\alpha_{uv}, p_v)} \tag{B.13}$$

$$C_t = (1 - \Theta_v) \underbrace{\sum_m^{t-1}(1 - k_\alpha\alpha_{uv})^m k_\alpha\alpha_{uv}}_{\phi'_{uv} = h(k_\alpha\alpha_{uv}, p_v)} \tag{B.14}$$

$$\phi'_{uv} > \phi_{uv} \; if \; k_\alpha > 1 \tag{B.15}$$

$$\phi'_{uv} < \phi_{uv} \; if \; k_\alpha < 1 \tag{B.16}$$

Observed period-specific $\widehat{VE}_t$ depends on the rate of testing (probability of testing if test-willing) p_τ, population share of test-willing individuals Θ_τ, and the ratio of testing probability amongst the vax-unwilling relative to the vax-willing, suppose test-willing k_τ.

$$\widehat{VE}_t = 1 - \frac{\widehat{A}_t}{\widehat{B}_t + \widehat{C}_t} \tag{B.17}$$

$$\widehat{A}_t = A_t \Theta_\tau \underbrace{\sum_m^t p_\tau \sum_n^{t-1}(1 - p_\tau)^n p_\tau \sum_q^{t-1}(1 - p_\tau)p_\tau^q \sum_r^{t-1}(1 - p_\tau)^{t-1-r}p_\tau^r}_{\Omega = j(p_\tau)} \tag{B.18}$$

$$\widehat{B}_t = B_t \Theta_\tau \Omega \tag{B.19}$$

$$\widehat{C}_t = C_t \Theta_\tau \underbrace{\sum_m^t k_\tau p_\tau \sum_n^{t-1}(1 - k_\tau p_\tau)^n k_\tau p_\tau \sum_q^{t-1}(1 - k_\tau p_\tau)k_\tau p_\tau^q \sum_r^{t-1}(1 - k_\tau p_\tau)^{t-1-r}(k_\tau p_\tau)^r}_{\Omega' = j(k_\tau p_\tau)} \tag{B.20}$$

$$\Omega' > \Omega \; if \; k_\tau > 1 \tag{B.21}$$

$$\Omega' < \Omega \; if \; k_\tau < 1 \tag{B.22}$$

$$\Omega' = \Omega \; if \; k_\tau = 1 \tag{B.23}$$

The bias of overall VE is then the difference between the weighted sums of the daily observed VEs and that of the daily true VEs.

$$VE = \sum_{t=0}^{T} w_t VE_t = \sum_{t=0}^{T} w_t \left(1 - \frac{\Theta_v \phi_v}{\Theta_v \phi_{uv} + (1 - \Theta_v)\phi'_{uv}} \right) \tag{B.24}$$

$$\widehat{VE} = \sum_{t=0}^{T} w_t \widehat{VE}_t = \sum_{t=0}^{T} w_t \left(1 - \frac{\Theta_v \phi_v \Theta_\tau \Omega}{\Theta_v \phi_{uv} \Theta_\tau \Omega + (1 - \Theta_v)\phi'_{uv} \Theta_\tau \Omega'} \right) \tag{B.25}$$

In this framework, the relative propensity of testing between the vax-unwilling and the vax-willing determines if observed $\widehat{VE}$ is unbiased. Independence between testing rate and vax-willingness ($k_\tau = 1$) yields zero theoretical bias ($Bias\{\hat{VE}\} = 0$).

$$Bias\{\widehat{VE}\} = VE - \widehat{VE} \tag{B.26}$$

$$Bias\{\widehat{VE}\} = \sum_{t=0}^{T} w_t \left(\frac{\Theta_v \phi_v}{\Theta_v \phi_{uv} + (1 - \Theta_v)\phi'_{uv}} - \frac{\Theta_v \phi_v \Theta_\tau \Omega}{\Theta_v \phi_{uv} \Theta_\tau \Omega + (1 - \Theta_v)\phi'_{uv} \Theta_\tau \Omega'} \right) \tag{B.27}$$

$$Bias\{\widehat{VE}\} = \sum_{t=0}^{T} w_t \left(\frac{\Theta_v \phi_v \left(\Theta_v \phi_{uv} \Theta_\tau \Omega + (1 - \Theta_v)\phi'_{uv} \Theta_\tau \Omega'\right) + \Theta_v \phi_v \Theta_\tau \Omega\left(\Theta_v \phi_{uv} + (1 - \Theta_v)\phi'_{uv}\right)}{\left(\Theta_v \phi_{uv} + (1 - \Theta_v)\phi'_{uv}\right)\left(\Theta_v \phi_{uv} \Theta_\tau \Omega + (1 - \Theta_v)\phi'_{uv} \Theta_\tau \Omega'\right)} \right) \tag{B.28}$$

$$Bias\{\widehat{VE}\} = \sum_{t=0}^{T} w_t \left(\frac{\Theta_v (1 - \Theta_v) \phi_v \phi'_{uv} \Theta_\tau (\Omega' - \Omega)}{\left(\Theta_v \phi_{uv} + (1 - \Theta_v)\phi'_{uv}\right)\left(\Theta_v \phi_{uv} \Theta_\tau \Omega + (1 - \Theta_v)\phi'_{uv} \Theta_\tau \Omega'\right)} \right) \tag{B.29}$$

$$Bias\{\widehat{VE}\} = 0 \; if \; k_\tau = 1 \tag{B.30}$$

All else held equal, theoretical bias is increasing in T (longer time period) but decreasing in p_v (probability of being vaccinated if vax-willing and not already vaccinated, which both ϕ_v and ϕ_{uv} are functions of). Theoretical bias also increases as k_τ (asymmetry in testing propensity between the vax-willing and vax-unwilling) and k_α (asymmetry in baseline attack rates between the vax-willing and vax-unwilling) diverge from 1.

w_t can be interpreted as the functional form of the choice of study design. All other parameters Ξ held equal, an unbiased study design solves for w_t, such that the dynamic VE equation is exactly the expectation of the VE distribution. The choice of w_t may result in

unbiased estimates of VE for particular sets of Ξ but not necessarily for some. An unbiased design with w_t, given parameters Ξ, further requires the observed $\hat{VE}$ equation to yield the expectation of the VE distribution.

$$w_t|\Xi \equiv \{\sum_{t=0}^{T} w_t VE_t = \mathbb{E}(VE)\} \tag{B.31}$$

$$w_t|\Xi \equiv \{\sum_{t=0}^{T} w_t \hat{VE}_t = \mathbb{E}(VE)\} \tag{B.32}$$

C   Scenario-Specific Parameter Assumptions

Scenario A entails slow vaccine rollout due to supply constraints arising from nascent manufacturing chains, that is limited to mostly trial settings. Comprehensive contact tracing and testing regimes were more common; together with a generally risk averse environment, imperfections in testing parameters are set to be limited. The vax-unwilling are further assumed to have higher baseline attack rates than the vax-willing due to disease-naivety and that this group is likely a selective category of disease deniers, e.g., that of COVID-19. This most likely reflects realities during the initial SARS-CoV-2 wave (Wild Type).

Scenario B assumes then that production of vaccines were scaled up, and testing became less comprehensive due to containment fatigue in the population, as well as stressed healthcare resources due to higher case load. This is more reflective of the end of 2020, when the Alpha, Beta, and Gamma variants became dominant.

Scenario C assumes that vaccines being rolled out at a faster pace, as major pharmaceutical companies were beginning to achieve economies of scale in production and orders. Due to even more extensive case load, testing became even patchier. Baseline attack rates are independent of attitudes towards vaccination. The Delta wave, which started in the second quarter of 2021, where lockdowns were imposed in response to Delta's perceived severity relative to earlier variants, making baseline risks symmetric between the vax-willing and vax-unwilling.

Scenario D assumes that vaccine production achieved economies of scale, and orders saturated, hence vaccine rollouts were rapid. Testing became more selective, comprehensive contact tracing and testing regimes are gradually abandoned. Baseline attack rates were also lower for the vax-unwilling than the vax-willing. This is reflective of the first Omicron wave at the end of 2021, contact tracing, testing, and self-regulatory behaviour on risky social

activities were gradually abandoned due to lower perceived severity of Omicron SARS-CoV-2 infections amid a high speed vaccine rollouts. A substantial fraction of countries were rolling out booster doses at this point. The vax-willing during this period may have a higher baseline risk due to a combination of repeated and undetected past infections amongst the unvaccinated, and vaccine mandates in place conditioning higher risk social activities on vaccination.

Scenario E further assumes that testing is even patchier as all forms of contact tracing, and mandated testing are abandoned. Moreover, vaccination rate is generally stable, as production lines have likely been optimised for efficiency. In the second Omicron wave at the start of 2022 as by this point, most countries have already completed their vaccination drives. Demand, rather than supply, is the constraint.

D   Drivers of Bias

Let $\delta_{M,\Xi}$ be the set of simulated bias for parameters Ξ, and design choice M. Further let $\delta_{\Xi}|M$ be the set of bias for the specific design M, and Ξ_M be the parameter set for that choice of design. We then express the vector of combined simulated biases $\delta_{M,\Xi}$ as a linear combination of design fixed effects λ_M, parameters Ξ, and a design-parameter-specific error term $\varepsilon_{M,\Xi}$. We may allow the bias-parameter nexus to vary across design choices, in lieu of a linear fixed effects term. Hence, we express the bias conditional on design M, $\delta_{\Xi}|M$ to be a linear combination of parameter set Ξ_M, and an error term $\varepsilon_{\Xi|M}$.

$$\delta_{M,\Xi} = \lambda_M + \Xi\lambda_{\Xi} + \varepsilon_{M,\Xi} \tag{D.1}$$

$$\delta_{\Xi}|M = \lambda_{0|M} + \Xi_M\lambda_{\Xi|M} + \varepsilon_{\Xi|M} \tag{D.2}$$

E   Additional Findings

E.1   Rational for Range of Colours in Heat Maps

The following charts show the contour maps of (1) the number needed to vaccinate (NNV) against VE, and (2) the first derivative of NNV against VE, both conditional on baseline attack rates between 1% and 11% at 2 ppt intervals. The algebraic derivation is as follows, where α_b is the baseline attack rate, and α_v the vaccinated attack rate.

$$NNV = \frac{1}{\alpha_b - \alpha_v} \tag{E.1}$$

$$VE = 100\left(1 - \frac{\alpha_v}{\alpha_b}\right) \implies \alpha_v = \alpha_b\left(1 - \frac{VE}{100}\right) \tag{E.2}$$

$$NNV = \frac{1}{\alpha_b - \alpha_b(1 - \frac{VE}{100})} = \frac{100}{\alpha_b VE} \tag{E.3}$$

$$\frac{\partial NNV}{\partial VE} = -\frac{100}{\alpha_b}VE^{-2} = -\frac{100}{\alpha_b VE^2} \tag{E.4}$$

There are three points to make here. Firstly, for any VE, NNV rises at an increasing rate as the baseline attack rate (α_b) decreases. Secondly, irrespective of α_b, NNV falls at a decreasing rate as VE rises. Thirdly, as α_b decreases, the rate of decline in NNV with respect to VE becomes more pronounced.

Figure E.1: Contour Map of the Number Needed to Vaccinate (NNV) Against Vaccine Effectiveness (VE), Conditional on Baseline Attack Rates
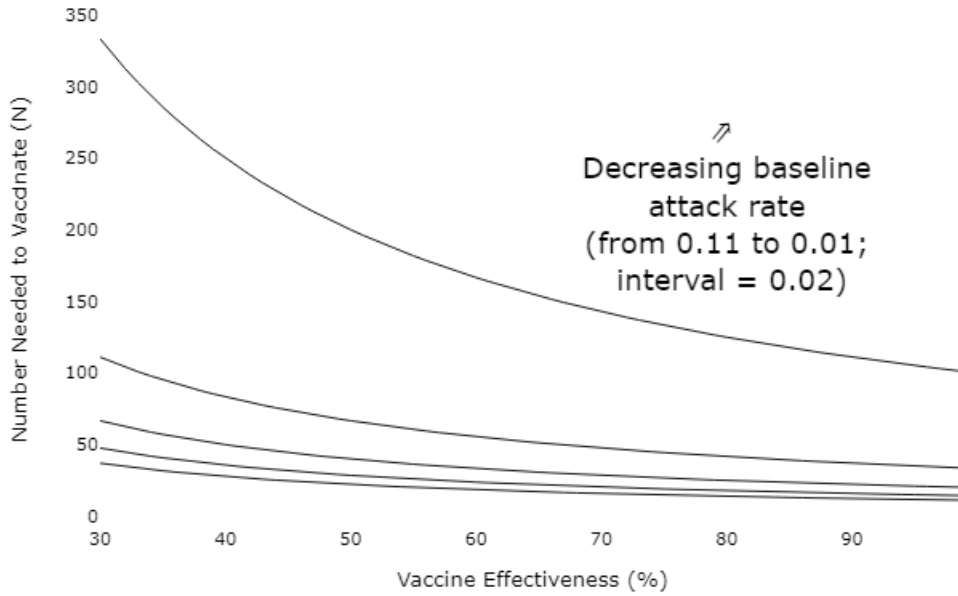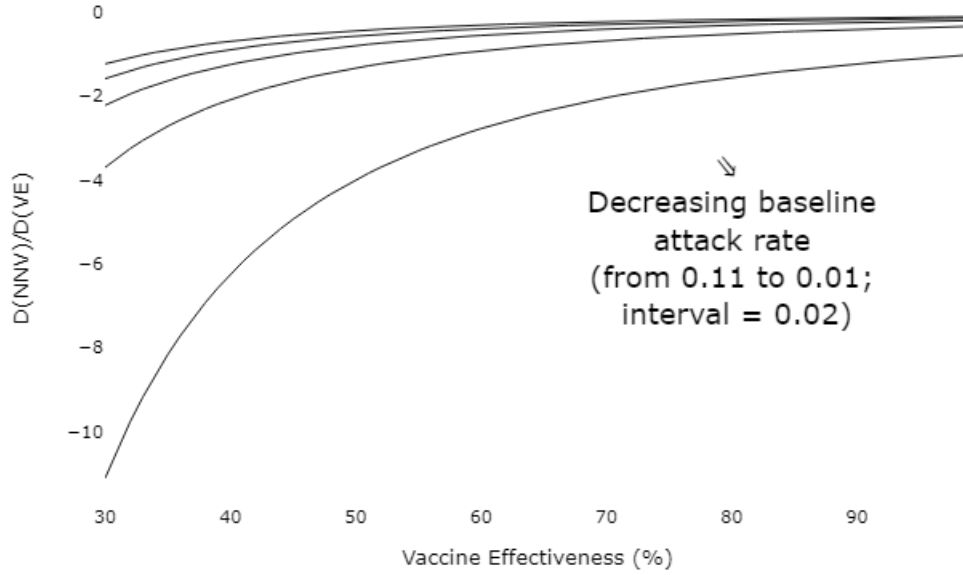
Figure E.2: Contour Map of the First Derivative of the Number Needed to Vaccinate (NNV) Against Vaccine Effectiveness (VE), Conditional on Baseline Attack Rates

E.2 Simulation

The following table contains a snippet of the full table of simulated bias across study designs for all 888125 parameter sets Ξ. The full document can be found here.

Figure E.3: Simulated bias in VE estimates across study designs

Index columns (constant for this segment): N = 0.5, T = 0.025. Remaining row labels: Design and mew_ve.

Design	mew_ve	Cohort (aggregate; N): Observed y	Cohort (aggregate; N): True y	Cohort (aggregate; PD): Observed y	Cohort (aggregate; PD): True y	Cohort (immortal time correction): Observed y	Cohort (immortal time correction): True y	Cohort: Observed y	Cohort: True y	TND: First Pos, First Neg	TND: First Pos, Multiple Neg
0.5	0.5	nan	-58.790000	-193.990000	-112.600000	-127.470000	-62.590000	-73.070000	-23.440000	-73.420000	-40.560000
	0.6	nan	-42.590000	-99.610000	-95.160000	-110.470000	-43.680000	-65.150000	-13.630000	-64.900000	-36.420000
	0.7	nan	-23.170000	-118.340000	-54.220000	-132.990000	-25.340000	-80.310000	-3.720000	-81.490000	-51.150000
	0.8	nan	0.620000	-76.300000	0.780000	-63.500000	-0.560000	-40.460000	7.630000	-40.300000	-23.550000
	0.9	nan	nan	-39.430000	-56.650000	-30.860000	-18.290000	-21.250000	-7.110000	-21.230000	-11.960000
0.75	0.5	nan	-22.590000	-152.910000	-44.240000	-76.330000	-20.030000	-41.350000	7.410000	-40.680000	-17.170000
	0.6	-54.300000	-15.300000	-74.160000	-38.210000	-61.730000	-12.500000	-34.230000	8.770000	-33.190000	-13.960000
	0.7	-68.630000	-5.470000	-87.670000	-17.520000	-76.560000	-4.380000	-46.710000	10.540000	-46.640000	-26.450000
	0.8	-38.450000	6.880000	-48.040000	9.260000	-37.460000	7.000000	-23.720000	12.760000	-23.170000	-11.500000
	0.9	-18.010000	nan	-22.620000	-34.890000	-16.180000	-7.700000	-11.090000	0.030000	-10.870000	-4.740000
0.9	0.5	nan	-12.040000	-137.390000	-32.110000	-65.300000	-9.330000	-35.010000	16.140000	-34.130000	-12.460000
	0.6	nan	nan	-65.430000	-28.220000	-52.660000	-4.520000	-29.070000	15.300000	-27.910000	-10.180000
	0.7	-58.250000	-0.180000	-76.930000	-11.430000	-66.270000	0.910000	-41.210000	14.610000	-40.940000	-22.370000
	0.8	-30.090000	8.670000	-39.130000	10.720000	-30.050000	8.920000	-19.030000	14.200000	-18.360000	-8.100000
	0.9	-13.460000	nan	-19.350000	-28.300000	-11.560000	-5.090000	-7.740000	1.760000	-7.460000	-2.410000
1.0	0.5	-57.100000	-8.540000	-133.940000	-23.450000	-58.230000	-4.970000	-29.460000	19.100000	-28.400000	-8.470000
	0.6	nan	nan	-63.490000	-20.710000	-46.670000	-1.330000	-24.560000	17.310000	-23.280000	-6.970000
	0.7	-52.090000	1.260000	-70.520000	-7.470000	-59.800000	2.870000	-36.420000	15.780000	-35.960000	-18.910000
	0.8	-26.560000	9.210000	-35.800000	nan	-26.780000	9.660000	-16.400000	14.640000	-15.680000	-6.240000
	0.9	-10.960000	nan	-15.320000	-26.270000	-9.220000	-4.010000	-5.830000	2.480000	-5.520000	-1.090000
1.1	0.5	-45.820000	nan	-106.120000	-16.600000	-43.260000	0.470000	-20.200000	22.880000	-18.850000	-1.470000
	0.6	-30.050000	nan	-47.830000	-15.160000	-34.550000	2.710000	-17.040000	20.090000	-15.570000	-1.350000
	0.7	-40.650000	3.810000	-49.240000	-3.720000	-44.510000	5.680000	-26.760000	17.600000	-25.950000	-11.600000
	0.8	-20.120000	10.170000	-23.460000	nan	-19.000000	10.760000	-11.110000	15.350000	-10.260000	-2.320000
	0.9	-7.450000	nan	-9.010000	-24.340000	-5.580000	-2.770000	-3.000000	3.330000	-2.630000	0.910000
1.25	0.5	-35.770000	nan	-95.400000	-9.480000	-34.040000	5.090000	-14.510000	26.550000	-12.980000	2.760000
	0.6	-24.500000	nan	-43.130000	-9.450000	-29.030000	6.020000	-13.880000	22.680000	-12.330000	0.950000
	0.7	-30.240000	6.100000	-40.520000	-0.130000	-33.910000	7.980000	-19.470000	19.330000	-18.380000	-6.320000
	0.8	-14.210000	11.020000	-19.080000	nan	-13.470000	11.630000	-7.110000	16.000000	-6.170000	0.520000

Image shows a segment of the full figure, which is available as a separate file with 888125 rows

An alternate version below normalises the simulated biases row-wise (by parameter sets), such that they are directly benchmarked against the survival analysis cohort without testing requisites. The full document can be found here.

Figure E.4: Relative bias in VE estimates across study designs; normalised to the survival analysis cohort without testing requisites

Index columns (constant for this segment): N = 0.5, T = 0.025. Remaining row labels: Design and mew_ve.

Design	mew_ve	Cohort (aggregate; N): Observed y	Cohort (aggregate; N): True y	Cohort (aggregate; PD): Observed y	Cohort (aggregate; PD): True y	Cohort (immortal time correction): Observed y	Cohort (immortal time correction): True y	Cohort: Observed y	Cohort: True y	TND: First Pos, First Neg	TND: First Pos, Multiple Neg
0.5	0.5	-0.000000	0.939287	3.099377	1.799009	2.036587	1.000000	1.167439	0.374501	1.173031	0.648027
	0.6	-0.000000	0.975046	2.280449	2.178571	2.529075	1.000000	1.491529	0.312042	1.485806	0.833791
	0.7	-0.000000	0.914365	4.670087	2.139700	5.248224	1.000000	3.169298	0.146803	3.215864	2.018548
	0.8	-0.000000	-1.107143	136.250000	-1.392857	113.392857	1.000000	72.250000	-13.625000	71.964286	42.053571
	0.9	-0.000000	-0.000000	2.155823	3.097321	1.687261	1.000000	1.161837	0.388737	1.160744	0.653909
0.75	0.5	-0.000000	1.127808	7.634049	2.208687	3.810784	1.000000	2.064403	-0.369945	2.030954	0.857214
	0.6	4.344000	1.224000	5.932800	3.056800	4.938400	1.000000	2.738400	-0.701600	2.655200	1.116800
	0.7	15.668950	1.248858	20.015982	4.000000	17.479452	1.000000	10.664384	-2.406393	10.648402	6.038813
	0.8	-5.492857	0.982857	-6.862857	1.322857	-5.351429	1.000000	-3.388571	1.322857	-3.310000	-1.642857
	0.9	2.333961	-0.000000	2.937662	4.531169	2.101299	1.000000	1.440260	-0.003896	1.411688	0.615584
0.9	0.5	-0.000000	1.290461	14.725616	3.441586	6.998928	1.000000	3.752412	-1.729904	3.658092	1.335477
	0.6	-0.000000	-0.000000	14.475664	6.243363	11.650442	1.000000	6.431416	-3.384956	6.174779	2.252212
	0.7	-64.010989	-0.197802	-84.538462	-12.560440	-72.824176	1.000000	-45.285714	16.054945	-44.989011	-24.582418
	0.8	-3.373318	0.971973	-4.386771	1.201794	-3.368834	1.000000	-2.133408	1.591928	-2.058296	-0.908072
	0.9	2.644401	-0.000000	3.801572	5.559921	2.271120	1.000000	1.520629	-0.345776	1.465619	0.473477
1.0	0.5	11.488934	1.718310	26.949698	4.718310	11.716298	1.000000	5.927565	-3.843058	5.714286	1.704225
	0.6	-0.000000	-0.000000	47.736842	15.571429	35.090226	1.000000	18.466165	-13.015038	17.503759	5.240602
	0.7	-18.149826	0.439024	-24.571429	-2.602787	-20.836237	1.000000	-12.689895	5.498258	-12.529617	-6.588850
	0.8	-2.749482	0.953416	-3.706004	0.000000	-2.772257	1.000000	-1.697723	1.515528	-1.623188	-0.645963
	0.9	2.733167	-0.000000	3.820449	6.551122	2.299252	1.000000	1.453865	-0.618454	1.376559	0.271820
1.1	0.5	-97.489362	0.000000	-225.787234	-35.319149	-92.042553	1.000000	-42.978723	48.680851	-40.106383	-3.127660
	0.6	-11.088561	0.000000	-17.649446	-5.594096	-12.749077	1.000000	-6.287823	7.413284	-5.745387	-0.498155
	0.7	-7.156690	0.670775	-8.669014	-0.654930	-7.836268	1.000000	-4.711268	3.098592	-4.568662	-2.042254
	0.8	-1.869888	0.945167	-2.180297	0.000000	-1.765799	1.000000	-1.032528	1.426380	-0.953532	-0.215613
	0.9	2.689531	-0.000000	3.252708	8.787004	2.014440	1.000000	1.083032	-1.202166	0.949458	-0.328520
1.25	0.5	-7.027505	0.000000	-18.742633	-1.862475	-6.687623	1.000000	-2.850688	5.216110	-2.550098	0.542240
	0.6	-4.069767	0.000000	-7.164452	-1.569767	-4.822259	1.000000	-2.305648	3.767442	-2.048173	0.157807
	0.7	-3.789474	0.764411	-5.077694	-0.016291	-4.249373	1.000000	-2.439850	2.422306	-2.303258	-0.791980
	0.8	-1.221840	0.947549	-1.640585	0.000000	-1.158212	1.000000	-0.611350	1.375752	-0.530525	0.044712

Image shows a segment of the full figure, which is available as a separate file with 888125 rows

The researcher may instead be interested in a direct comparison of the study designs in terms of absolute bias. The version below first converts all estimation biases into absolutes, and then normalise to that of the survival analysis cohort without testing requisites. The full document can be found here.
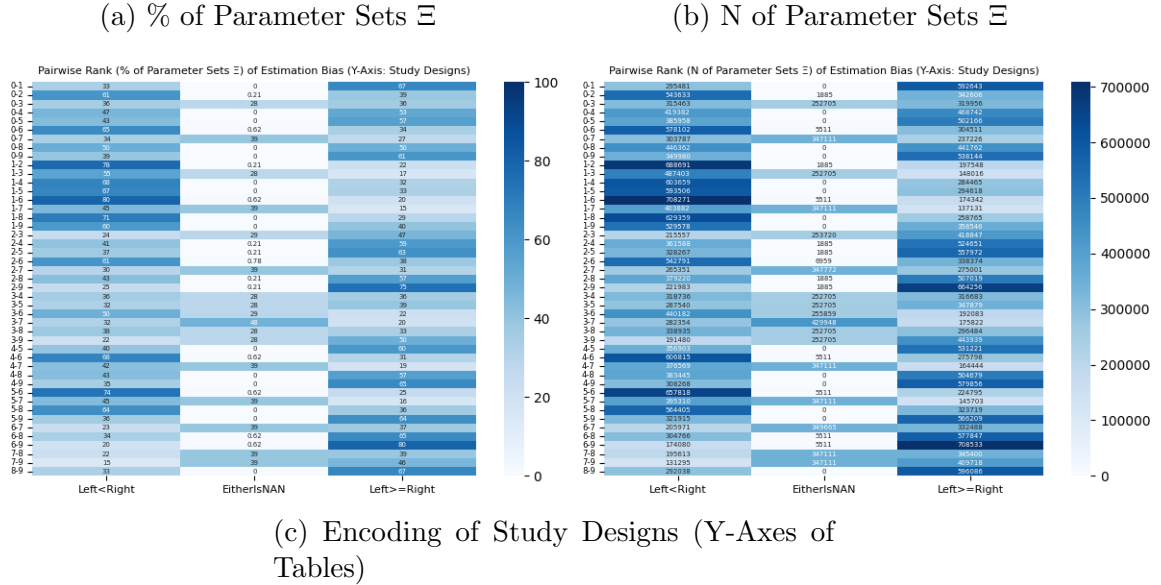
Figure E.5: Relative absolute bias in VE estimates across study designs; normalised to the survival analysis cohort without testing requisites

		Design	Cohort (aggregate; N): Observed y	Cohort (aggregate; N): True y	Cohort (aggregate; PD): Observed y	Cohort (aggregate; PD): True y	Cohort (immortal time correction): Observed y	Cohort (immortal time correction): True y	Cohort: Observed y	Cohort: True y	TND: First Pos, First Neg	TND: First Pos, Multiple Neg
N T Theta_v p_v Theta_tau p_tau ktau alpha_b kalpha mew_ve												
	0.5	0.5	0.000000	0.939287	3.099377	1.799009	2.036587	1.000000	1.167439	0.374501	1.173031	0.648027
		0.6	0.000000	0.975046	2.280449	2.178571	2.529075	1.000000	1.491529	0.312042	1.485806	0.833791
		0.7	0.000000	0.914365	4.670087	2.139700	5.248224	1.000000	3.169298	0.146803	3.215864	2.018548
		0.8	0.000000	1.107143	136.250000	1.392857	113.392857	1.000000	72.250000	13.625000	71.964286	42.053571
		0.9	0.000000	0.000000	2.155823	3.097321	1.687261	1.000000	1.161837	0.388737	1.160744	0.653909
	0.75	0.5	0.000000	1.127808	7.634049	2.208687	3.810784	1.000000	2.064403	0.369945	2.030954	0.857214
		0.6	4.344000	1.224000	5.932800	3.056800	4.938400	1.000000	2.738400	0.701600	2.655200	1.116800
		0.7	15.668950	1.248858	20.015982	4.000000	17.479452	1.000000	10.664384	2.406393	10.648402	6.038813
		0.8	5.492857	0.982857	6.862857	1.322857	5.351429	1.000000	3.388571	1.822857	3.310000	1.642857
		0.9	2.338961	0.000000	2.937662	4.531169	2.101299	1.000000	1.440260	0.003896	1.411688	0.615584
	0.9	0.5	0.000000	1.290461	14.725616	3.441586	6.998928	1.000000	3.752412	1.729904	3.658092	1.335477
		0.6	0.000000	0.000000	14.475664	6.243363	11.650442	1.000000	6.431416	3.384956	6.174779	2.252212
		0.7	64.010989	0.197802	84.538462	12.560440	72.824176	1.000000	45.285714	16.054945	44.989011	24.582418
		0.8	3.373318	0.971973	4.386771	1.201794	3.368834	1.000000	2.133408	1.591928	2.058296	0.908072
		0.9	2.644401	0.000000	3.801572	5.559921	2.271120	1.000000	1.520629	0.345776	1.465619	0.473477
0.5 0.025	1.0	0.5	11.488934	1.718310	26.949698	4.718310	11.716298	1.000000	5.927565	3.843058	5.714286	1.704225
		0.6	0.000000	0.000000	47.736842	15.571429	35.090226	1.000000	18.466165	13.015038	17.503759	5.240602
		0.7	18.149826	0.439024	24.571429	2.602787	20.836237	1.000000	12.689895	5.498258	12.529617	6.588850
		0.8	2.749482	0.953416	3.706004	0.000000	2.772257	1.000000	1.697723	1.515528	1.623188	0.645963
		0.9	2.733167	0.000000	3.820449	6.551122	2.299252	1.000000	1.453865	0.618454	1.376559	0.271820
	1.1	0.5	97.489362	0.000000	225.787234	35.319149	92.042553	1.000000	42.978723	48.680851	40.106383	3.127660
		0.6	11.088561	0.000000	17.649446	5.594096	12.749077	1.000000	6.287823	7.413284	5.745387	0.498155
		0.7	7.156690	0.670775	8.669014	0.654930	7.836268	1.000000	4.711268	3.098592	4.568662	2.042254
		0.8	1.869888	0.945167	2.180297	0.000000	1.765799	1.000000	1.032528	1.426580	0.953532	0.215613
		0.9	2.689531	0.000000	3.252708	8.787004	2.014440	1.000000	1.083032	1.202166	0.949458	0.328520
	1.25	0.5	7.027505	0.000000	18.742633	1.862475	6.687623	1.000000	2.850688	5.216110	2.550098	0.542240
		0.6	4.069767	0.000000	7.164452	1.569767	4.822259	1.000000	2.305648	3.767442	2.048173	0.157807
		0.7	3.789474	0.764411	5.077694	0.016291	4.249373	1.000000	2.439850	2.422306	2.303258	0.791980
		0.8	1.221840	0.947549	1.640585	0.000000	1.158212	1.000000	0.611350	1.375752	0.530525	0.044712

Image shows a segment of the full figure, which is available as a separate file with 888125 rows

In the following figure, subfigure (a) shows the % of parameter sets for which one study design yields higher VE than another, vice versa. Subfigure (b) shows the equivalent in number of parameter sets. This pairwise comparison of all 10 study designs across parameter sets suggest that rank reversals are common. However, even though there is no study design that maintained pairwise superiority for more than 80% of parameter sets, these rankings do not account for the 'realism' or real-world frequency of the corresponding parameter settings. Hence, when weighted against the real-world occurrence of respective parameter sets (e.g., settings reflective of Omicron as in scenarios D and E may be more likely than others), rank reversals may be infrequent, and that dominant study designs may exist.

Figure E.6: Pairwise Rank of Estimation Bias (Y-Axis: Study Designs)

(a) % of Parameter Sets Ξ

(b) N of Parameter Sets Ξ

(c) Encoding of Study Designs (Y-Axes of Tables)

|	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

E.3 Drivers of Bias

When rescaled to parameter values in scenario D, bias is less severe in most study designs, except those highly affected by dynamic vaccine coverage (p_v), and testing imperfections (Θ_τ, p_τ, k_τ). Reflective of parameter choice ($k_\alpha = 0.75$), deviation from symmetric baseline attack rates between the vax-willing and vax-unwilling may bias VE estimates only by a small but visible degree. The testing rate amongst the test-willing (p_τ), asymmetry in testing rates between the vax-willing and vax-unwilling (k_τ), and vaccination rate (p_v) still exert meaningful bias on VE estimates. Importantly, as these sensitivity estimates are homothetic, the TNDs remain most robust to 'less-than-ideal' conditions than the cohorts.

Figure E.7: Rescaled Association Between Parameters Ξ and Absolute Bias By Study Design $(p_v = 0.5; \Theta_\tau = 0.75; p_\tau = 0.5; k_\tau = 0.75; k_\alpha = 0.75)$

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

(i)   (j)
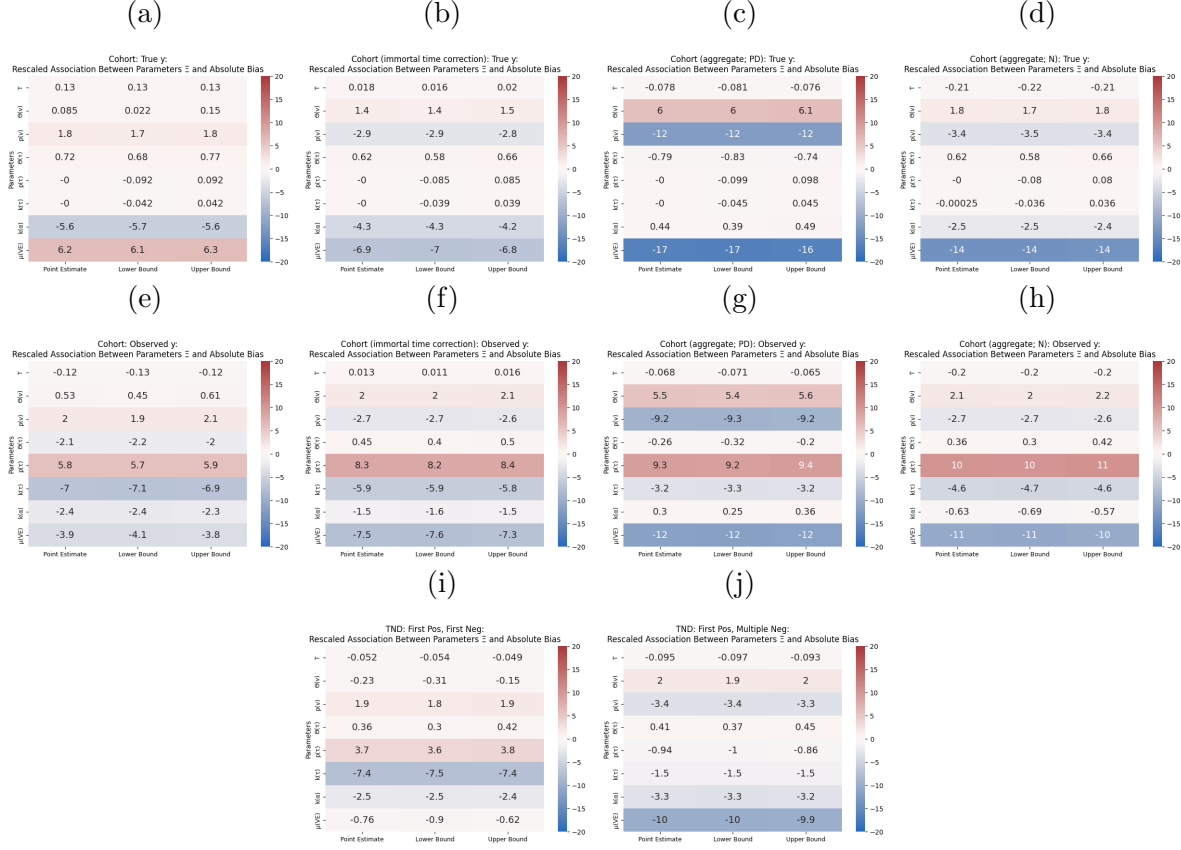
An alternate set of parameter values entails higher testing rate $(p_\tau = 0.75)$, and almost symmetric testing rate between the vax-willing and vax-unwilling $(k_\tau = 0.95)$. As the degree of bias arising from testing frictions almost disappeared, barring study designs whose estimation bias are highly sensitive to vaccination rates (p_v), and asymmetry in baseline attack rates (k_α), estimation bias can be small if confounding is absent, and if residual sources of bias are minimal, especially both TNDs. There are immense potential gains for policymakers if steps are taken to mitigate selective testing in respective infectious disease surveillance systems, such as through randomness in community sampling.

Figure E.8: Rescaled Association Between Parameters Ξ and Absolute Bias By Study Design $(p_v = 0.5;\ \Theta_\tau = 0.75;\ p_\tau = 0.75;\ k_\tau = 0.95;\ k_\alpha = 0.75)$

(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)
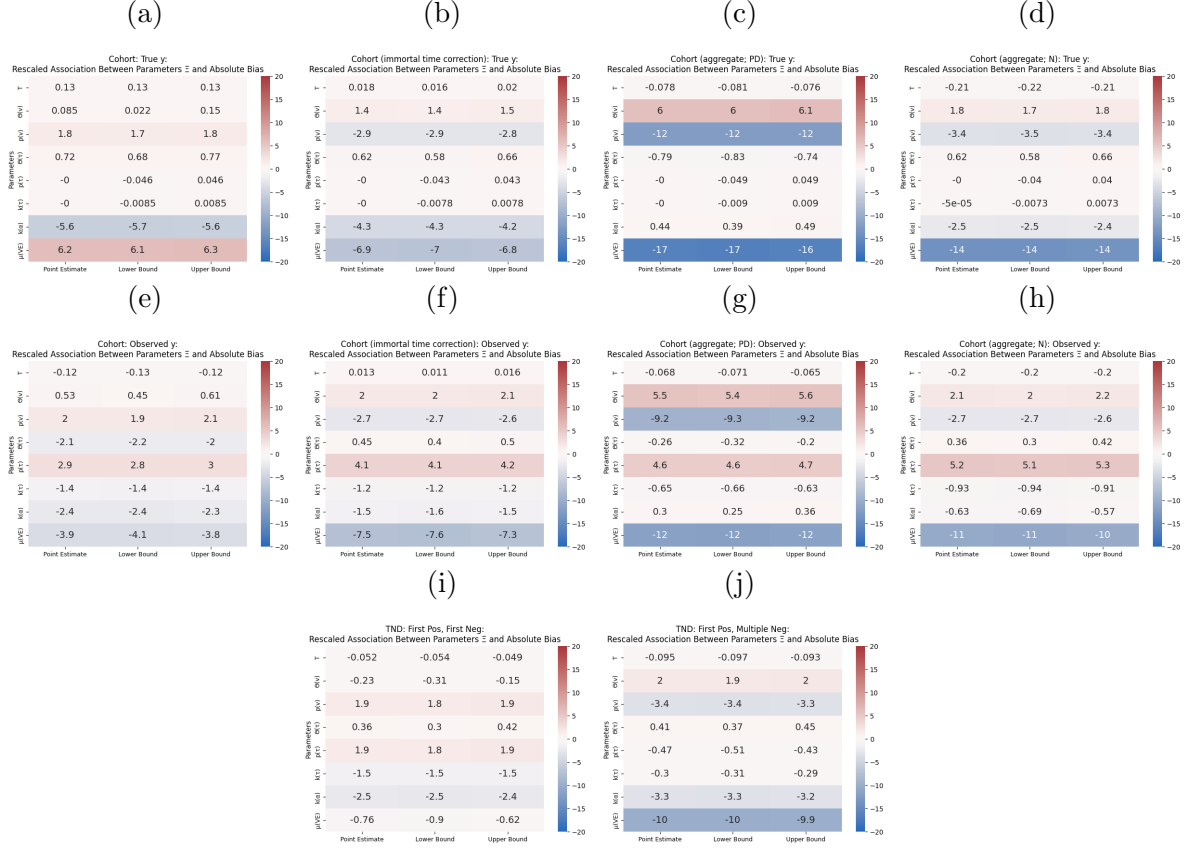
As a summative exercise, the FE model indicates that all factors except for follow-up duration (T), and deviation from full test-willing population (Θ_τ), are important within-design drivers of estimation bias on average.

Figure E.9: Association Between Parameters Ξ and Absolute Bias