

**Defining and Estimating Outcomes Directly Averted by a Vaccination Program when  
Rollout Occurs Over Time**

**Authors:** Katherine M Jia<sup>1§</sup>, Christopher B Boyer<sup>2,3</sup>, Alyssa Bilinski<sup>4,5</sup>, Marc Lipsitch<sup>1,6</sup>

**Affiliations:**

<sup>1</sup> Center for Communicable Disease Dynamics, Department of Epidemiology,  
Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

<sup>2</sup> Department of Quantitative Health Sciences, Cleveland Clinic Research, Cleveland, Ohio, USA

<sup>3</sup> Department of Medicine, Cleveland Clinic Lerner College of Medicine of Case Western Reserve  
University, Cleveland, Ohio, USA

<sup>4</sup> Department of Health Services, Policy, and Practice, Brown University School of Public Health,  
Providence, Rhode Island, USA

<sup>5</sup> Department of Biostatistics, Brown University School of Public Health, Providence, Rhode  
Island, USA

<sup>6</sup> Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health,  
Boston, Massachusetts, USA

**§ Correspondence:** Katherine M Jia, Ph.D., Center for Communicable Disease Dynamics,  
Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington  
Avenue, Boston, MA 02115, USA. +1 (617) 432-7955. (kjia@hsph.harvard.edu).

**Sources of funding:** K.M.J and M.L. were supported by a subaward from CDC cooperative  
agreement to Carnegie-Mellon University (Grant number: U01 IP001121). A.B. was supported by  
National Institutes of General Medical Studies (Grant number: 1R35GM155224).

## Abstract (248 words)

During the COVID-19 pandemic, estimating the total deaths averted by vaccination has been of great public health interest. Instead of estimating total deaths averted by vaccination among both vaccinated and unvaccinated individuals, some studies empirically estimated only “directly averted” deaths among vaccinated individuals, typically suggesting that vaccines prevented more deaths overall than directly due to the indirect effect. Here, we define the causal estimand to quantify outcomes “directly averted” by vaccination—i.e., the impact of vaccination for vaccinated individuals, holding vaccination coverage fixed—for vaccination at multiple time points, and show that this estimand is a lower bound on the total outcomes averted when the indirect effect is non-negative. We develop an unbiased estimator for the causal estimand in a one-stage randomized controlled trial (RCT) and explore the bias of a popular “hazard difference” estimator frequently used in empirical studies. We show that even in an RCT, the hazard difference estimator is biased if vaccination has a non-null effect, as it fails to incorporate the greater depletion of susceptibles among the unvaccinated individuals. In simulations, the overestimation is small for averted deaths when infection-fatality rate is low, as for many important pathogens. However, the overestimation can be large for averted infections given a high basic reproduction number. Additionally, we define and compare estimand and estimators for avertible outcomes (i.e., outcomes that could have been averted by vaccination, but were not due to failure to vaccinate). Future studies can explore the identifiability of the causal estimand in observational settings.

**Keywords:** vaccine-averted outcomes, cumulative incidence difference, direct impact, overall impact

## 1 Introduction

During the COVID-19 pandemic, determining the total number of infections (or deaths) averted by vaccination has been of great public health interest.<sup>1–10</sup> Researchers are interested in how many infections (or deaths) have been averted overall by COVID-19 vaccine rollout programs, compared to the counterfactual of no vaccination for anyone. However, in the presence of indirect effects, the key challenge is that we may not observe a comparable population that is unvaccinated throughout. Rather than estimating the total number of outcomes averted among both vaccinated and unvaccinated individuals, some empirical studies<sup>4–8</sup> have instead estimated outcomes “directly averted” among vaccinated individuals by conditioning on the actual vaccination coverage in the rest of the population. The estimand and estimation procedures used in selected examples of studies are summarized in **eAppendix 1**. Typically, studies that estimated outcomes directly averted among vaccinated individuals computed daily or weekly hazards of death among all vaccinated and not-yet-vaccinated individuals, calculated the difference in hazards, multiplied this difference with the number of vaccinated survivors, and summed the results across time. We refer to this method as the “hazard difference estimator.” Empirical analyses of this type commonly assumed that such directly averted outcomes are a lower bound on the total averted outcomes due to the indirect effect in reducing transmission.<sup>4–6</sup>

This study is motivated by two research gaps from the numerous empirical analyses<sup>4–8,11,12</sup> that estimated directly averted (or avertible) outcomes among vaccinated (or unvaccinated) individuals under a vaccine rollout. First, the causal estimand for directly averted outcomes has not been precisely defined as a mathematical quantity for vaccination at multiple time points under interference. Second, the popular hazard difference estimator used by these analyses has not been evaluated. Therefore, we first propose the causal estimand and its unbiased estimator based on a

one-stage randomized controlled trial (RCT). We use the causal estimand to formalize the lower bound assumption and identify the condition under which a vaccine rollout program has averted more outcomes overall than directly among the vaccinated individuals. Last, we evaluate the bias of the hazard difference estimator relative to the causal estimand. This paper is an important extension to our previous study<sup>13</sup> which develops estimands for quantifying averted outcomes for vaccination at a single time point, as vaccination almost always occurs over time in reality.

Section 2 of this paper describes the setup and notation. Section 3 defines the causal estimand for quantifying directly averted outcomes among vaccinated individuals, as well as its unbiased estimator and the hazard difference estimator. Section 4 examines the bias of the hazard difference estimator both analytically and through simulations.

## 2 Setup and notation

Hudgens and Halloran<sup>14</sup> defined four *effect* estimands (namely the direct, indirect, total, and overall effects) for vaccination at a single time point in a two-stage randomized trial, which reduces to a one-stage RCT when the study population consists of a single group. In reality, two-stage RCTs are rarely conducted or justified; for our purpose of estimating directly averted outcomes, we define our estimand and estimators based on a one-stage RCT, with notation closely aligned to that of Hudgens and Halloran.<sup>14</sup>

Consider a one-stage RCT that consists of  $N$  individuals indexed by  $j = 1, \dots, N$  with a large  $N$ .<sup>14</sup> Consider  $q + 2$  evenly spaced measurement intervals for  $q \in \mathbb{N}$ . Let  $l \in \{0, \dots, q + 1\}$  denote each interval with baseline measurements taken in interval 0, and  $q + 1$  representing the end of follow-up.<sup>15</sup> Let  $X_j$  denote the assigned vaccination time, where  $x_j \in \{0, \dots, q + 1\}$  are possible realizations of  $X_j$ .<sup>16</sup> Vaccination occurs *at the beginning* of each interval, whereas  $x_j = q + 1$  denotes unvaccinated throughout. Let  $\mathbf{X} = (X_1, \dots, X_N)$  denote the vaccination times individuals

were assigned. Let  $\mathbf{x}$  denote a possible realization of  $\mathbf{X}$ . Let  $\mathcal{X}(N)$  denote the set of all possible  $(q + 1)^N$  vaccination time allocations for the group, for which  $\mathbf{x} \in \mathcal{X}(N)$ . Throughout we assume perfect compliance (i.e., assignment to a particular vaccination time is equivalent to receipt of vaccination at that time if the person is still alive), no loss-to-follow-up, and no measurement error. In an ideal one-stage RCT, these assumptions are expected to hold.

Here, interference is assumed—that is, the potential outcome for any individual depend on vaccination assignments of every other individual in the group.<sup>14</sup> Let  $Y_{q+1,j}(\mathbf{x}) \in \{0,1\}$  be a *cumulative* indicator of experiencing the outcome (e.g., death) *before the beginning* of interval  $q + 1$  for individual  $j$  had the group followed the vaccination schedule  $\mathbf{x} \in \mathcal{X}(N)$ . By convention,  $Y_{0,j} \equiv 0$ .<sup>17</sup>

Let  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$  denote parameterizations that govern the distribution of  $\mathbf{X}$ , where  $q + 1$  is the number of potential vaccination times,  $d$  is the number of days of each interval,  $\rho_x$  is the proportion of individuals assigned to  $x \in \{0, \dots, q + 1\}$  such that  $\sum_{x=0}^{q+1} \rho_x = 1$ . We assume  $\boldsymbol{\rho}$  is a *mixed individual assignment strategy*,<sup>14,18</sup> as defined in **eAppendix 2**. In words,  $\boldsymbol{\rho}$  randomly assigns  $\rho_0 \times 100\%$  of the individuals to receive vaccination at baseline,  $\rho_1 \times 100\%$  to receive vaccination at the beginning of interval 1, and so on, with  $\rho_{q+1} \times 100\%$  to remain unvaccinated throughout.<sup>18</sup> Let  $\boldsymbol{\Phi} = \{q + 1, d; \mathbf{0}\}$  denote no vaccination. To quantify vaccine-averted outcomes, our goal is to assess the impact of some vaccination strategy  $\boldsymbol{\rho}$  compared to  $\boldsymbol{\Phi}$ .

At baseline, individuals are randomly assigned to  $\mathbf{X}$  conditional on a mixed individual assignment strategy (i.e., fixed proportions of individuals [e.g., 20%, 30%, 50%] were assigned to receive vaccination at specific times [e.g., Day 0, Day 60, Day 120], respectively; see definition in **eAppendix 2**). Note that random assignments  $\mathbf{X}$  occurs only at baseline, even though individuals receive vaccination at different times. This study design is referred to as a “one-stage” RCT

because randomization takes place only once, in contrast to the “two-stage” RCT described by Hudgens and Halloran,<sup>14</sup> in which groups are first randomized to different strategies and then individuals within each group are randomized to vaccination conditional on their group’s assigned strategy. Let  $\bar{Y}_{q+1}(x; \mathbf{p})$  denote the group average potential outcome as defined in **eAppendix 2**, which is equivalent to population average potential outcome<sup>14</sup> because there is only one group. Let  $\Delta\bar{Y}_{l+1}(x; \mathbf{p}) = \bar{Y}_{l+1}(x; \mathbf{p}) - \bar{Y}_l(x; \mathbf{p})$  for  $l \in \{0, \dots, q\}$  be the difference in  $\bar{Y}$  between intervals  $l$  and  $l + 1$ .<sup>19</sup>

### 3 Causal estimand and estimators for quantifying outcomes directly averted by vaccination

#### 3.1 Causal estimand

In our prior work with vaccination at a single time point,<sup>13</sup> we defined the “direct impact” estimand to quantify outcomes directly averted among vaccinated individuals as the number of vaccinated individuals multiplied by the direct effect (DE)<sup>14</sup> for vaccination at the baseline time point. In notation, the number of outcomes directly averted by  $\mathbf{p} = \{1, d; \rho_0\}$  (i.e, vaccination of  $\rho_0$  at the beginning of a single interval, with interval duration  $d$ ), compared to no vaccination  $\mathbf{\Phi} = \{1, d; 0\}$ , is:

$$\begin{aligned}\delta_1^D(\mathbf{\Phi}, \mathbf{p}) &= N\rho_0(\bar{Y}_1(1; \mathbf{p}) - \bar{Y}_1(0; \mathbf{p})) \\ &= N\rho_0 DE_1((1, 0); \mathbf{p}).\end{aligned}\quad (1)$$

When vaccination occurs at multiple time points, there could be multiple versions of direct effects. A direct effect can be a contrast between non-vaccination and some vaccination time  $x' \in \{0, \dots, q\}$ , conditional on  $\mathbf{p}$  (e.g., comparing individuals unvaccinated versus vaccinated at interval 0, or unvaccinated versus vaccinated at interval 1). In notation, the direct effect comparing probability of having developed the outcome by the beginning of interval  $q + 1$  for an individual

unvaccinated throughout versus assigned to  $x'$  when the group follows strategy  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$  is:

$$DE_{q+1}((q + 1, x'); \boldsymbol{\rho}) = \bar{Y}_{q+1}(q + 1; \boldsymbol{\rho}) - \bar{Y}_{q+1}(x'; \boldsymbol{\rho}). \quad (2)$$

Now, extending the estimand in equation (1) for vaccination at two time points (i.e.,  $x' = 0$  or  $x' = 1$ , as compared to unvaccinated  $x = 2$ ), we consider the direct effect of  $x' \in \{0,1\}$  compared to unvaccinated throughout, weighted by the number of individuals assigned with  $x'$ , and summed across  $x' \in \{0,1\}$ . In notation, the number of outcomes directly averted by  $\boldsymbol{\rho} = \{2, d; \rho_0, \rho_1\}$ , compared to no vaccination  $\boldsymbol{\Phi} = \{2, d; \mathbf{0}\}$ , is:

*Definition 1 (Causal estimand for directly averted outcomes for vaccination at two time points)*

$$\begin{aligned} \delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) &= N \cdot \{\rho_0 \cdot [\bar{Y}_2(2; \boldsymbol{\rho}) - \bar{Y}_2(0; \boldsymbol{\rho})] + \rho_1 \cdot [\bar{Y}_2(2; \boldsymbol{\rho}) - \bar{Y}_2(1; \boldsymbol{\rho})]\} \\ &= N \cdot \{\rho_0 \cdot DE_2((2,0); \boldsymbol{\rho}) + \rho_1 \cdot DE_2((2,1); \boldsymbol{\rho})\}. \end{aligned} \quad (3)$$

**eAppendix 3** extends  $\delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$  to an arbitrary number of vaccination times. **eAppendix 4** shows that  $\delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$  is the lower bound on outcomes averted among both vaccinated and unvaccinated individuals when the indirect effect is non-negative. However, the indirect effect can be negative when transmission or fatality parameters vary over time, as we showed earlier for vaccination at a single time point.<sup>13</sup> Therefore, in the more general case of vaccination at multiple time points, the indirect effect is not guaranteed to be non-negative under many realistic scenarios.

### 3.2 Unbiased estimator

Define  $\hat{Y}_2(x; \boldsymbol{\rho}) = \frac{\sum_{j=1}^N \gamma_{2,j}(\mathbf{X}) I[X_j=x]}{\sum_{j=1}^N I[X_j=x]}$  for  $x \in \{0,1,2\}$ . That is,  $\hat{Y}_2(x; \boldsymbol{\rho})$  is the cumulative incidence by the beginning of interval 2 for individuals assigned to  $x$  under strategy  $\boldsymbol{\rho}$ . Then the estimator for  $\delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$  in equation (3) is:

$$\widehat{\delta}_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) = N \cdot \{\rho_0 \cdot [\hat{Y}_2(2; \boldsymbol{\rho}) - \hat{Y}_2(0; \boldsymbol{\rho})] + \rho_1 \cdot [\hat{Y}_2(2; \boldsymbol{\rho}) - \hat{Y}_2(1; \boldsymbol{\rho})]\}. \quad (4)$$

**eAppendix 5** proves that  $\widehat{\delta_{q+1}^D}(\boldsymbol{\phi}, \boldsymbol{\rho})$  is an unbiased estimator for  $\delta_{q+1}^D(\boldsymbol{\phi}, \boldsymbol{\rho})$  for  $q \in \mathbb{N}$  in a one-stage RCT under mixed assignment strategy  $\boldsymbol{\rho}$ . Note that  $\rho_x$  for  $x \in \{0, 1, \dots, q+1\}$  is known and fixed under a mixed assignment strategy  $\boldsymbol{\rho}$  in an RCT, although it must be estimated in an observational setting.

### 3.3 Hazard difference estimator

As summarized in the literature review (eAppendix 1), recent empirical studies<sup>4–8</sup> often used what we refer to as the *hazard difference* estimator, which relies on the number at risk and the number of new cases among individuals vaccinated (and not-yet-vaccinated) by a given time to quantify outcomes directly averted among vaccinated individuals.

Let  $\Delta\hat{Y}_{l+1}(x; \boldsymbol{\rho}) = \hat{Y}_{l+1}(x; \boldsymbol{\rho}) - \hat{Y}_l(x; \boldsymbol{\rho})$  for  $l \in \{0, \dots, q\}$ . Equation (5) defines the hazard difference estimator for two vaccination times (see eAppendix 6 for extension to an arbitrary number of vaccination times).

*Definition 2 (Hazard difference estimator for directly averted outcomes for vaccination at two time points)*

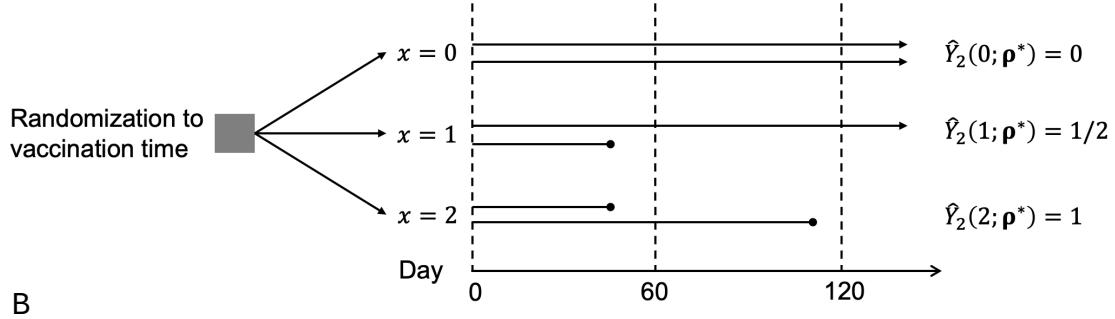
$$\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho}) = \hat{N}_0^v(\boldsymbol{\rho}) \left( \hat{h}_1^u(\boldsymbol{\rho}) - \hat{h}_1^v(\boldsymbol{\rho}) \right) + \hat{N}_1^v(\boldsymbol{\rho}) \left( \hat{h}_2^u(\boldsymbol{\rho}) - \hat{h}_2^v(\boldsymbol{\rho}) \right) \quad (5)$$

where  $\hat{h}_1^v(\boldsymbol{\rho}) = \Delta\hat{Y}_1(0; \boldsymbol{\rho})$  is the incidence among vaccinated individuals by the beginning of interval 1,  $\hat{h}_1^u(\boldsymbol{\rho}) = \frac{\rho_1 \Delta\hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta\hat{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2}$  is the incidence among not-yet-vaccinated individuals by the beginning of interval 1 (Note this quantity is the combination of two distinct groups—individuals assigned to  $x = 1$  and  $x = 2$ ),  $\hat{N}_0^v(\boldsymbol{\rho}) = N\rho_0$  is the number survived by the beginning of interval 0 among individuals assigned to  $x = 0$ ,  $\hat{N}_1^v(\boldsymbol{\rho}) = N \left[ \rho_0 \left( 1 - \hat{Y}_1(0; \boldsymbol{\rho}) \right) + \rho_1 \left( 1 - \hat{Y}_1(1; \boldsymbol{\rho}) \right) \right]$  is the combined number survived by the beginning of interval 1 among individuals

assigned to  $x = 0$  and  $x = 1$ ,  $\hat{h}_2^v(\boldsymbol{\rho}) = \frac{\rho_0 \Delta \hat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \hat{Y}_2(1; \boldsymbol{\rho})}{\rho_0(1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1(1 - \hat{Y}_1(1; \boldsymbol{\rho}))}$  is the hazard among vaccinated individuals by the beginning of interval 2,  $\hat{h}_2^u(\boldsymbol{\rho}) = \frac{\Delta \hat{Y}_2(2; \boldsymbol{\rho})}{1 - \hat{Y}_1(2; \boldsymbol{\rho})}$  is the hazard among not-yet-vaccinated individuals by the beginning of interval 2.

To give intuition for the unbiased estimator and the hazard difference estimator, consider the following example (**Figure 1**). Suppose we observe an ideal trial randomizing 6 individuals into three arms according to the strategy  $\boldsymbol{\rho}^* = \{2, 60; \frac{1}{3}, \frac{1}{3}\}$ . One-third of the individuals are randomly assigned to vaccination at baseline ( $x = 0$ ), one-third to vaccination at the beginning of interval 1 ( $x = 1$ ), and the remaining to no vaccination throughout ( $x = 2$ ). As illustrated in **Figure 1A**, the unbiased estimator gives  $\widehat{\delta}_2^D(\boldsymbol{\phi}, \boldsymbol{\rho}^*) = 6 \cdot \{\rho_0 \cdot [\hat{Y}_2(2; \boldsymbol{\rho}^*) - \hat{Y}_2(0; \boldsymbol{\rho}^*)] + \rho_1 \cdot [\hat{Y}_2(2; \boldsymbol{\rho}^*) - \hat{Y}_2(1; \boldsymbol{\rho}^*)]\} = 6 \cdot \left[\frac{1}{3} \cdot (1 - 0) + \frac{1}{3} \cdot \left(1 - \frac{1}{2}\right)\right] = 3$ . As illustrated in **Figure 1B**, the hazard difference estimator gives  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho}^*) = \hat{N}_0^v(\boldsymbol{\rho}^*) \left(\hat{h}_1^u(\boldsymbol{\rho}^*) - \hat{h}_1^v(\boldsymbol{\rho}^*)\right) + \hat{N}_1^v(\boldsymbol{\rho}^*) \left(\hat{h}_2^u(\boldsymbol{\rho}^*) - \hat{h}_2^v(\boldsymbol{\rho}^*)\right) = 2 \cdot \left(\frac{1}{2} - 0\right) + 3 \cdot (1 - 0) = 4$ .

A



**FIGURE 1** | Schematic representation of an ideal randomized controlled trial with 6 individuals (horizontal lines) under the strategy  $\mathbf{p}^* = \{2, 60; \frac{1}{3}, \frac{1}{3}\}$ . One-third of individuals are randomly assigned to vaccination at baseline ( $x = 0$ ), one third to vaccination at the beginning of interval 1 ( $x = 1$ ), and the remaining to no vaccination throughout ( $x = 2$ ). The dashed vertical lines represent the beginning of each interval. (A) The unbiased estimator considers the cumulative incidence in each arm. Note  $\hat{Y}_l(x; \mathbf{p}^*)$  is the cumulative incidence by the beginning of interval  $l$  among individuals assigned to vaccination time  $x$  under strategy  $\mathbf{p}^*$ . (B) Hazard difference estimator considers the hazards and survival among vaccinated and not-yet-vaccinated individuals. Note  $\hat{h}_{l+1}^v(\cdot)$  (or  $\hat{h}_{l+1}^u(\cdot)$ ) is the hazard for vaccinated (or not-yet-vaccinated) individuals from interval  $l$  to  $l + 1$ .

#### 4 Bias of the hazard difference estimator for the causal estimand

Now, we use analytical and simulation approaches to examine the bias of  $\widehat{\delta}_2^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho})$  relative to the causal estimand.

#### 4.1 Analytic comparison

First, re-write the causal estimand  $\delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$  as follows (See derivation in **eAppendix 3**):

$$\begin{aligned} \delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) &= N \cdot \left\{ \rho_0 \cdot \left( \frac{\rho_1 \Delta \bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta \bar{Y}_1(0; \boldsymbol{\rho}) \right) \right. \\ &\quad \left. + [(\rho_0 + \rho_1) \cdot \Delta \bar{Y}_2(2; \boldsymbol{\rho}) - (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho}))] \right\} \quad (6) \end{aligned}$$

Then, expand  $E \left[ \widehat{\delta}_2^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho}) \right]$  as follows (See derivation in equation [S13] **eAppendix 6**):

$$\begin{aligned} E \left[ \widehat{\delta}_2^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] &= E \left[ \widehat{N}_0^v(\boldsymbol{\rho}) (\widehat{h}_1^u(\boldsymbol{\rho}) - \widehat{h}_1^v(\boldsymbol{\rho})) + \widehat{N}_1^v(\boldsymbol{\rho}) (\widehat{h}_2^u(\boldsymbol{\rho}) - \widehat{h}_2^v(\boldsymbol{\rho})) \right] \\ &= N \left\{ \rho_0 \cdot \left( \frac{\rho_1 \Delta \bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta \bar{Y}_1(0; \boldsymbol{\rho}) \right) \right. \\ &\quad + E \left[ (\rho_0 + \rho_1) \cdot \frac{\rho_0 (1 - \widehat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \widehat{Y}_1(1; \boldsymbol{\rho}))}{(\rho_0 + \rho_1) (1 - \widehat{Y}_1(2; \boldsymbol{\rho}))} \cdot \Delta \widehat{Y}_2(2; \boldsymbol{\rho}) \right] \\ &\quad \left. - (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho})) \right\}. \quad (7) \end{aligned}$$

If  $\frac{\rho_0 (1 - \widehat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \widehat{Y}_1(1; \boldsymbol{\rho}))}{(\rho_0 + \rho_1) (1 - \widehat{Y}_1(2; \boldsymbol{\rho}))} = 1$ , then equation (7) equals (6) (See proof in **eAppendix 6**). That

is,  $\widehat{\delta}_2^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho})$  is an unbiased estimator under the null (i.e., when the survival is the same between those assigned to no vaccination and those assigned with  $x = 0$  or  $1$ ). However, if

$\frac{\rho_0(1-\hat{Y}_1(0;\rho))+\rho_1(1-\hat{Y}_1(1;\rho))}{(\rho_0+\rho_1)(1-\hat{Y}_1(2;\rho))} \neq 1$ , then  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \rho)$  is biased relative to the causal estimand, implying

that it is biased if vaccination has a non-null effect.

**Appendix 7** provides an alternative unbiased estimator for quantifying directly averted outcomes with a similar (but not identical) expression and have same data requirement as  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \rho)$ . The alternative unbiased estimator allows estimation of directly averted deaths using data aggregated by vaccination status.

## 4.2 Simulation comparison

### 4.2.1 Scenarios

We simulate an epidemic with strategy  $\boldsymbol{\rho}' = \{2, 60; 0.2, 0.3\}$  under different infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ), or vaccine efficacy against death given infection ( $VE_{death}$ ). We examine the bias of the hazard difference estimator  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho}')$  relative to the causal estimand  $\delta_2^D(\boldsymbol{\phi}, \boldsymbol{\rho}')$ , where  $\boldsymbol{\phi} = \{2, 60; \mathbf{0}\}$ , and identify the conditions under which the bias would be substantial (**Table 1**).

**TABLE 1** | Scenarios for simulations, varied by infection-fatality rate, vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ )

Scenario	Infection-fatality rate	Vaccine efficacy
Scenario 1	1%	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario 2	10%	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario 3	100%	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario 4	1%	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario 5	10%	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario 6	100%	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario 7	1%	$VE_{inf} = 90\%; VE_{death} = 90\%$
Scenario 8	10%	$VE_{inf} = 90\%; VE_{death} = 90\%$
Scenario 9	100%	$VE_{inf} = 90\%; VE_{death} = 90\%$

In the main text, we explore scenarios with varying IFR,  $VE_{inf}$ , and  $VE_{death}$ , including several extreme scenarios for illustrative purposes. For sensitivity analyses, we explore scenarios with varying number of effective contacts ( $\beta$ ), as well as more realistic parameter values for  $\beta$ , IFR,

$VE_{inf}$ , and  $VE_{death}$ , specifically corresponding to seasonal flu, measles, and COVID-19 (wild-type strain).

#### 4.2.2 Model

Consider a hypothetical RCT with strategy  $\rho' = \{2, 60; 0.2, 0.3\}$ , in which interval 0 spans Days 0-59, interval 1 spans Days 60-119, and interval 2 is the post-follow-up period on or after Day 120, such that 20% of individuals can be assigned to vaccination at Day 0, 30% to vaccination at Day 60, or the rest remain unvaccinated throughout. In the susceptible-infected-recovered-death (SIRD) model, individuals are stratified by time of vaccination. Within each stratum, we specify a continuous-time SIRD model. The subscript 0 represents those assigned to receive vaccination in the beginning of Day 0, 1 for those assigned to receive vaccination in the beginning of Day 60, and 2 for the never vaccinated. For those receiving vaccination in the beginning of Day 60, vaccine efficacies against infection ( $\theta_1$ ) and death given infection ( $\kappa_1$ ) are time-varying variables that come in effect on and after Day 60 (**Table S4**). The SIRD model is defined in term of continuous time  $t$  as follows:

$$\left. \begin{aligned}
\frac{dS_2(t)}{dt} &= -\lambda(t) \cdot S_2(t) \\
\frac{dS_1(t)}{dt} &= -\theta_1(t) \cdot \lambda(t) \cdot S_1(t) \\
\frac{dS_0(t)}{dt} &= -\theta \cdot \lambda(t) \cdot S_0(t) \\
\frac{dI_2(t)}{dt} &= \lambda(t) \cdot S_2(t) - \gamma \cdot I_2(t) \\
\frac{dI_1(t)}{dt} &= \theta_1(t) \cdot \lambda(t) \cdot S_1(t) - \gamma \cdot I_1(t) \\
\frac{dI_0(t)}{dt} &= \theta \cdot \lambda(t) \cdot S_0(t) - \gamma \cdot I_0(t) \\
\frac{dR_2(t)}{dt} &= (1 - \mu) \cdot \gamma \cdot I_2(t) \\
\frac{dR_1(t)}{dt} &= (1 - \kappa_1(t) \cdot \mu) \cdot \gamma \cdot I_1(t) \\
\frac{dR_0(t)}{dt} &= (1 - \kappa \cdot \mu) \cdot \gamma \cdot I_0(t) \\
\frac{dD_2(t)}{dt} &= \mu \cdot \gamma \cdot I_2(t) \\
\frac{dD_1(t)}{dt} &= \kappa_1(t) \cdot \mu \cdot \gamma \cdot I_1(t) \\
\frac{dD_0(t)}{dt} &= \kappa \cdot \mu \cdot \gamma \cdot I_0(t)
\end{aligned} \right\} \quad (8)$$

where  $\gamma$  = recovery rate,  $\lambda(t) = \beta \cdot \frac{I_0(t) + I_1(t) + I_2(t)}{N(t)}$  (the hazard rate of infection),  $\beta$  = the number

of effective contacts made by a typical infectious individual per unit time,  $\mu$  = probability of death due to infection,  $N(t)$  = sum of all individuals alive at  $t$ . **eAppendix 8** shows the parameters and initial values used in simulations. For simulation, the model was discretized to day time-steps. Code is available at [https://github.com/katjia/vax\\_rollout\\_impact](https://github.com/katjia/vax_rollout_impact).

#### 4.2.3. Simulations

**Figure 2** compares the expected value of hazard difference estimator (i.e.,  $E \left[ \widehat{\delta}_2^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho}') \right]$ ) with the causal estimand (i.e.,  $\delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho}')$ ) under Scenarios outlined in **Table 1**; **eAppendix 9** shows the

bias of the hazard difference estimator relative to the causal estimand on the absolute and relative scales.

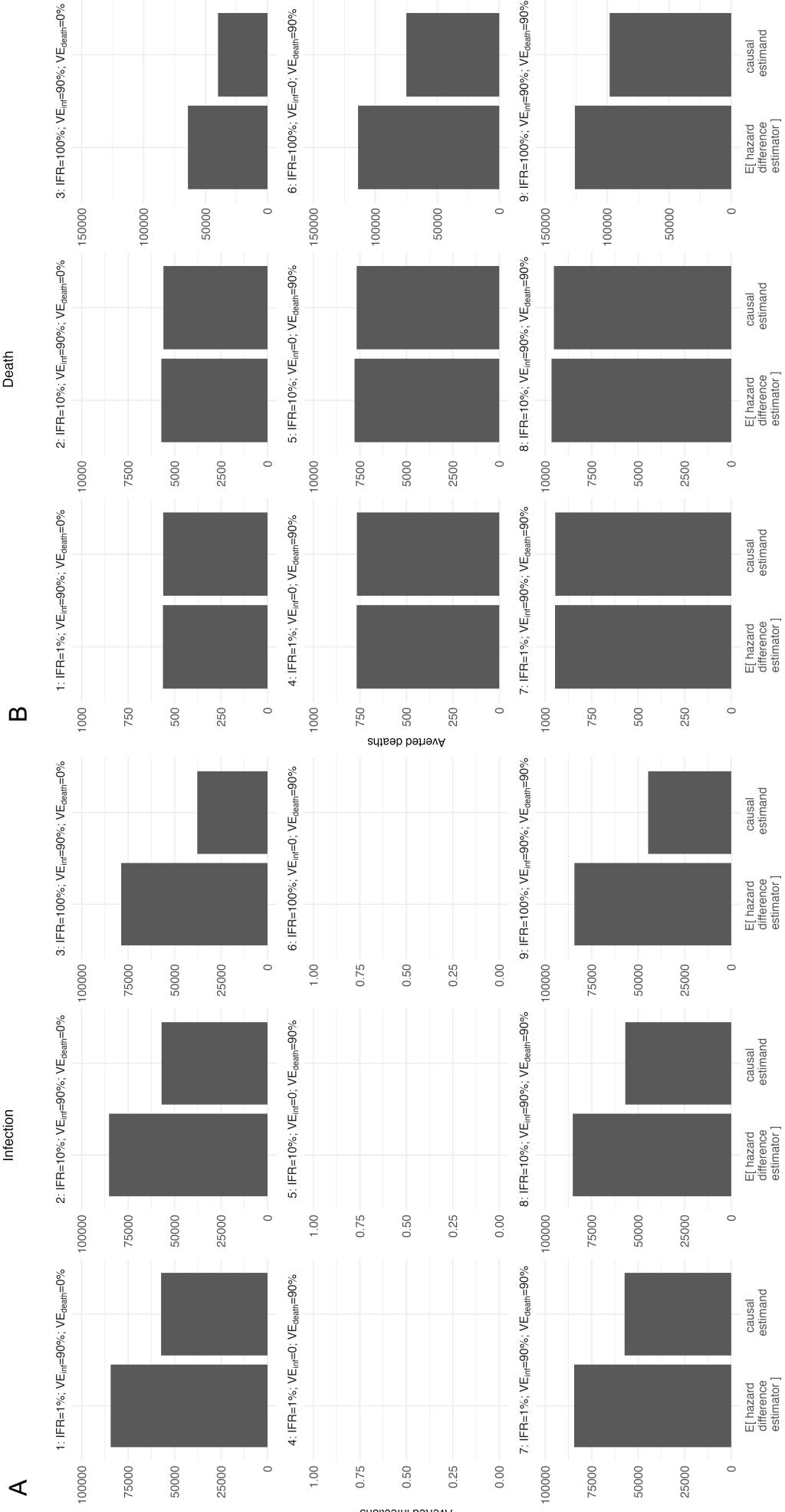
No infections were averted when  $VE_{inf} = 0\%$  (**Figure 2A**; Scenarios 4 to 6). In Scenarios where  $VE_{inf} = 90\%$ , the hazard difference estimator substantially overestimates the averted infections (**Figure 2A**; Scenarios 1 to 3, 7 to 9). This is because given the high reproduction number ( $R_0 \approx 3.57$ ) and a high  $VE_{inf}$ , susceptibles were preferentially depleted among the unvaccinated individuals quickly, such that by interval 1, individuals with  $x = 2$  had lower survival from infection than the average survival among individuals with  $x = 0$  and  $x = 1$  (**Figure S3**)—that is,  $\frac{\rho'_0(1-\hat{Y}_1(0;\boldsymbol{\rho}'))+\rho'_1(1-\hat{Y}_1(1;\boldsymbol{\rho}'))}{(\rho'_0+\rho'_1)(1-\hat{Y}_1(2;\boldsymbol{\rho}'))} > 1$  in equation (7). **eAppendix 11** varies the Scenarios by the number of effective contacts ( $\beta$ ) and show that the overestimation is more pronounced the higher  $\beta$  is.

The hazard difference estimator substantially overestimates the averted deaths when  $IFR=100\%$ , while the overestimation is trivial when  $IFR \leq 10\%$  (**Figure 2B**). This is due to the similar survival from death between vaccinated and unvaccinated individuals when  $IFR$  is low—

that is,  $\frac{\rho'_0(1-\hat{Y}_1(0;\boldsymbol{\rho}'))+\rho'_1(1-\hat{Y}_1(1;\boldsymbol{\rho}'))}{(\rho'_0+\rho'_1)(1-\hat{Y}_1(2;\boldsymbol{\rho}'))} \approx 1$ .

**eAppendix 6** repeats the same analyses for vaccination at additional time points and over a longer period, which shows that averted infections are even more severely overestimated compared with vaccination at two time points. Consistent with the main-text finding, the hazard difference estimator slightly overestimates averted deaths when  $IFR \leq 10\%$ . In **eAppendix 12**, we consider more realistic parameter values for seasonal flu, measles, and COVID-19 (wild-type strain). As with the main analyses,  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho}')$  substantially overestimates averted infections for measles due to high basic reproduction number ( $R_0 = 18$ ) and high  $VE_{inf}$ . It also slightly overestimates averted infections for seasonal flu and COVID-19 (wild-type), given a low  $R_0$  (i.e.,  $\leq 2.2$ ) and  $VE_{inf} > 0$ .

The overestimation of averted deaths is trivial due to the low IFR ( $\leq 3\%$  for all pathogens) (**Figure S5**).



**FIGURE 2 |** Infections (A) and deaths (B) directly averted by strategy  $\rho' = \{2,60; 0,2,0,3\}$  under different scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection (VE<sub>inf</sub>), and vaccine efficacy against death given infection (VE<sub>death</sub>).

## 6 Discussion

Recent empirical studies have estimated COVID-19 outcomes directly averted by vaccine rollout programs among vaccinated individuals. Here, we define a causal estimand for quantifying directly averted outcomes for vaccination at multiple time points and develop an unbiased estimator. We also examine a popular estimator used by recent empirical studies—the hazard difference estimator (as we call it)—and showed that it is biased relative to the causal estimand when vaccination has a non-null effect, as it fails to incorporate the preferential depletion of susceptibles among the unvaccinated individuals. The simulations performed here, albeit limited, suggest that the bias is substantial for averted infections, as susceptibles were preferentially depleted among the unvaccinated group quickly (due to high effective contacts and high vaccine efficacy against infection). On the other hand, the bias for averted deaths is small when IFR is  $\leq 10\%$  (since survival is similar between vaccinated and unvaccinated individuals), as is the case for many important infections.

Empirical studies frequently used the hazard difference estimator. As a measure, the hazard among (un)vaccinated individuals is restricted to those who have not experienced the outcome between baseline and the start of interval  $k$  for  $k \geq 1$ . Consequently, the interval-specific hazard is subject to differential depletion of susceptibles among the unvaccinated group over time (assuming vaccines have protective effects).<sup>20</sup> However, the hazard difference estimator multiplies the number of vaccinated survivors by the hazard difference between vaccinated and unvaccinated individuals—implicitly assuming that the unvaccinated individuals have the same counterfactual survival as the vaccinated individuals, thereby failing to account for such differential depletion of susceptibles among the unvaccinated group. As our simulations show, the hazard difference estimator overestimates the number of infections averted when susceptibles are preferentially and

quickly depleted among unvaccinated individuals given a high reproduction number ( $R_0 \approx 3.57$ ) and a high  $VE_{inf}$ . The bias is less pronounced for averted deaths when IFR is low, as survival between vaccinated and unvaccinated individuals is similar. Therefore, for COVID-19 studies using the hazard difference estimator, the averted infection estimate is likely more biased than the averted death estimate.

Researchers can use data by aggregated by vaccination status to estimate averted outcomes by following the procedures outlined in **eAppendices 7 and 13**. Most publicly available data from vaccine registries are aggregated by vaccination status due to privacy or other considerations. In these situations, we recommend clearly specifying the intervals to which survival and hazard pertain. An example dataset is shown in **Table S11**, which organizes the hazard and the number of survivors within the same row, specifying the hazard to be one interval ahead of the number of survivors. This arrangement facilitates applying formulas for the alternative unbiased estimator or hazard difference estimator. The table also includes an explanatory footnote clarifying the intervals corresponding to the hazard and the number of survivors.

In the main text, we focus on vaccine-averted outcomes. Some other empirical studies estimated vaccine-avertible deaths—deaths that could have been averted by vaccination, but were not because of a failure to vaccinate. They used the hazard difference estimator for outcomes directly *avertible* by vaccination by multiplying the hazard difference with the number of unvaccinated survivors and summing across weeks.<sup>11,12</sup> **eAppendix 14** defines the estimand, the unbiased estimator, and the hazard difference estimator for outcomes directly avertible, and shows simulation results under the same Scenarios as outlined in **Table 1**. Compared to the causal estimand, the hazard difference estimator could recover a similar value for avertible deaths when IFR is modest ( $\leq 10\%$ ) or when vaccines are highly effective at preventing death given infection

( $VE_{\text{death}} = 90\%$ ) (**Figures S6 & S7**). Note that when estimating directly avertible outcomes (or when comparing to any other counterfactual vaccination strategy than no vaccination), researchers need data disaggregated by vaccination time to use the unbiased estimator.

One major limitation is that the causal estimand and estimators proposed here are developed in the context of an ideal RCT assuming no confounding, no selection bias, perfect compliance, and no other sources of biases. In **eAppendix 13**, we discussed identifying the causal estimand for averted outcomes using observational data aggregated by vaccination status in the absence of confounding. In most observational settings, however, there may be strong confounding by vaccine uptake (i.e., individuals who choose to be vaccinated are also more likely to avoid infection, or elderly individuals are more likely to be vaccinated and also more likely to die from infection). In addition, individuals may reduce protective behaviors after vaccination (i.e., risk compensation),<sup>23</sup> and those who have been infected are also less likely to receive vaccination.<sup>22</sup> Confounding is also a concern in existing studies that estimate averted deaths from observational data, as they adjust for it by stratifying on only simplistic covariates (**Table S1**). Although observational studies violate the assumptions of an ideal RCT, addressing all these violations is beyond the scope of this paper. Future research could use observational data to emulate a target trial similar to the one described here for estimating vaccine-averted outcomes.

In conclusion, motivated by recent empirical studies estimating outcomes directly averted by vaccine rollout programs, we define a causal estimand for directly averted outcomes under interference, which is a lower bound on the total outcomes averted in the entire population when indirect effect is non-negative. We also develop an unbiased estimator in the context of a one-stage RCT and examine the bias of a popular estimator (the hazard difference estimator). The hazard difference estimator is biased relative to the causal estimand when vaccination has a non-null effect

because it does not incorporate differential depletion of susceptibles among unvaccinated individuals. Our simulations, albeit limited, show that the hazard difference estimator could substantially overestimate the averted infections when the basic reproduction number and vaccine efficacy against infection are high, while the overestimation is small for averted deaths under modest IFR, which is the case for many important infections.

## References

1. Vilches TN, Moghadas SM, Sah P, et al. Estimating COVID-19 Infections, Hospitalizations, and Deaths Following the US Vaccination Campaigns During the Pandemic. *JAMA Netw Open*. 2022;5(1):e2142725. doi:10.1001/jamanetworkopen.2021.42725
2. Schneider EC, Shah A, Sah P, et al. Impact of U.S. COVID-19 Vaccination Efforts: An Update on Averted Deaths, Hospitalizations, and Health Care Costs Through March 2022. To the Point. April 8, 2022. Accessed December 10, 2023. <https://doi.org/10.26099/d3dm-fa91>
3. Gavish N, Yaari R, Huppert A, Katriel G. Population-level implications of the Israeli booster campaign to curtail COVID-19 resurgence. *Sci Transl Med*. 2022;14(647):eabn9836. doi:10.1126/scitranslmed.abn9836
4. Haas EJ, McLaughlin JM, Khan F, et al. Infections, hospitalisations, and deaths averted via a nationwide vaccination campaign using the Pfizer–BioNTech BNT162b2 mRNA COVID-19 vaccine in Israel: a retrospective surveillance study. *The Lancet Infectious Diseases*. 2022;22(3):357-366. doi:10.1016/S1473-3099(21)00566-1
5. Brault A, Hart A, Uribe P, et al. Direct impact of COVID-19 vaccination in Chile: averted cases, hospitalizations, ICU admissions, and deaths. *BMC Infect Dis*. 2024;24(1):467. doi:10.1186/s12879-024-09304-1
6. Santos CVBD, Noronha TGD, Werneck GL, Struchiner CJ, Villela DAM. Estimated COVID-19 severe cases and deaths averted in the first year of the vaccination campaign in Brazil: a retrospective observational study. *The Lancet Regional Health - Americas*. 2023;17:100418. doi:10.1016/j.lana.2022.100418
7. Kayano T, Sasanami M, Kobayashi T, et al. Number of averted COVID-19 cases and deaths attributable to reduced risk in vaccinated individuals in Japan. *The Lancet Regional Health - Western Pacific*. 2022;28:100571. doi:10.1016/j.lanwpc.2022.100571
8. Kayano T, Nishiura H. Assessing the COVID-19 vaccination program during the Omicron variant (B.1.1.529) epidemic in early 2022, Tokyo. *BMC Infect Dis*. 2023;23(1):748. doi:10.1186/s12879-023-08748-1
9. Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*. 2022;22(9):1293-1302. doi:10.1016/S1473-3099(22)00320-6
10. Shoukat A, Vilches TN, Moghadas SM, et al. Lives saved and hospitalizations averted by COVID-19 vaccination in New York City: a modeling study. *The Lancet Regional Health - Americas*. 2022;5:100085. doi:10.1016/j.lana.2021.100085
11. Jia KM, Hanage WP, Lipsitch M, et al. Estimated preventable COVID-19-associated deaths due to non-vaccination in the United States. *Eur J Epidemiol*. Published online April 24, 2023. doi:10.1007/s10654-023-01006-3

12. Zhong M, Glazer T, Kshirsagar M, et al. Estimating Vaccine-Preventable COVID-19 Deaths Among Adults Under Counterfactual Vaccination Scenarios in The United States: A Modeling Study Using Observational Data. *J Pharm Pharmacol Res.* 2023;07(03). doi:10.26502/fjppr.079
13. Jia KM, Boyer CB, Wallinga J, Lipsitch M. Causal Estimands for Analyses of Averted and Avertible Outcomes due to Infectious Disease Interventions. *Epidemiology.* 2025;36(3):363-373. doi:10.1097/EDE.0000000000001839
14. Hudgens MG, Halloran ME. Toward Causal Inference With Interference. *Journal of the American Statistical Association.* 2008;103(482):832-842. doi:10.1198/016214508000000292
15. Young JG, Logan RW, Robins JM, Hernán MA. Inverse Probability Weighted Estimation of Risk Under Representative Interventions in Observational Studies. *Journal of the American Statistical Association.* 2019;114(526):938-947. doi:10.1080/01621459.2018.1469993
16. Boyer C, Lipsitch M. Emulating target trials of postexposure vaccines using observational data. *American Journal of Epidemiology.* doi:10.1093/aje/kwae350
17. Shi J, Swanson SA, Kraft P, Rosner B, De Vivo I, Hernán MA. Instrumental variable estimation for a time-varying treatment and a time-to-event outcome via structural nested cumulative failure time models. *BMC Med Res Methodol.* 2021;21(1):258. doi:10.1186/s12874-021-01449-w
18. Sobel ME. What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference. *Journal of the American Statistical Association.* 2006;101(476):1398-1407. doi:10.1198/016214506000000636
19. Janvin M, Stensrud MJ. Quantification of Vaccine Waning as a Challenge Effect. *Journal of the American Statistical Association.* Published online December 6, 2024:1-11. doi:10.1080/01621459.2024.2408776
20. Hernán MA. The Hazards of Hazard Ratios. *Epidemiology.* 2010;21(1):13-15. doi:10.1097/EDE.0b013e3181c1ea43
21. Brewer NT, Cuite CL, Herrington JE, Weinstein ND. Risk compensation and vaccination: Can getting vaccinated cause people to engage in risky behaviors? *ann behav med.* 2007;34(1):95-99. doi:10.1007/BF02879925
22. Hulme WJ, Williamson E, Horne EMF, et al. Challenges in Estimating the Effectiveness of COVID-19 Vaccination Using Observational Data. *Ann Intern Med.* 2023;176(5):685-693. doi:10.7326/M21-4269

**eAppendix of “Defining and Estimating Outcomes Directly Averted by a Vaccination Program when Rollout Occurs Over Time”**

**Table of Contents**

<b>eAppendix 1. Literature review of studies estimating the number of COVID-19 outcomes directly averted among vaccinated individuals or the total number of outcomes averted among both vaccinated and unvaccinated individuals .....</b>	<b>3</b>
<b>eAppendix 2. Defining individual, group, and population average potential outcomes.....</b>	<b>7</b>
<b>eAppendix 3. Causal estimand for outcomes directly averted by vaccination.....</b>	<b>9</b>
<b>eAppendix 4 Conditions under which directly averted outcomes is a lower bound on overall averted outcomes.....</b>	<b>12</b>
<b>eAppendix 5. Unbiased estimator for the causal estimand in a one-stage randomized controlled trial.....</b>	<b>17</b>
<b>eAppendix 6. Extending the hazard difference estimator for vaccination at more than two time points.....</b>	<b>20</b>
<b>eAppendix 7. An alternative estimator for quantifying directly averted outcomes using RCT data aggregated by vaccination status.....</b>	<b>27</b>
<b>eAppendix 8. Parameters for simulations in the main text.....</b>	<b>29</b>
<b>eAppendix 9. Bias of the hazard difference estimator relative to the causal estimand on the absolute and relative scales .....</b>	<b>31</b>
<b>eAppendix 10. Disease dynamics under each scenario .....</b>	<b>32</b>
<b>eAppendix 11. Scenarios for simulations by varying the number of effective contacts .....</b>	<b>33</b>

<b>eAppendix 12. Simulations with more realistic parameter values .....</b>	<b>36</b>
<b>eAppendix 13. Using observational data aggregated by vaccination status for estimation .</b>	<b>40</b>
<b>eAppendix 14. Outcomes directly avertible by vaccination .....</b>	<b>43</b>
<b>References.....</b>	<b>55</b>

**eAppendix 1. Literature review of studies estimating the number of COVID-19 outcomes directly averted among vaccinated individuals or the total number of outcomes averted among both vaccinated and unvaccinated individuals**

Recent studies estimated COVID-19 outcomes averted by vaccine rollout programs either by targeting the outcomes directly averted among vaccinated individuals<sup>1-5</sup> or by targeting the total outcomes averted among both vaccinated and unvaccinated individuals.<sup>6-10</sup> To better understand the estimation methodologies, we reviewed selected literature listed in **Table S1**.

TABLE S1 | Summary of selected studies that estimated COVID-19 outcomes averted by vaccine rollout programs

Year	Authors	Country	Title	COVID-19 Outcomes	Estimand <sup>a</sup>	Estimation methods
2021	Haas et al <sup>1</sup>	Israel	Infections, hospitalizations, and deaths averted via a nationwide vaccination campaign using the Pfizer–BioNTech BNT162b2 mRNA COVID-19 vaccine in Israel: a retrospective surveillance study	Infections, hospitalizations, severe or critical hospitalizations, deaths	Direct impact	<ol style="list-style-type: none"> <li>For each age group and each day, calculate the number of vaccinated individuals at risk of developing the outcome by multiplying the number at risk with the proportion vaccinated. Also calculate the number of unvaccinated individuals by subtracting the numbers of vaccinated individuals from the total population at risk.</li> <li>Calculate the hazard rate difference between unvaccinated and vaccinated individuals.</li> <li>Estimate the daily averted outcomes by multiplying the number of vaccinated individuals at risk with the hazard rate difference.</li> <li>Sum the daily averted outcomes across days within each age group, and then sum across age groups.</li> </ol>
2024	Brault et al <sup>2</sup>	Chile	Direct impact of COVID-19 vaccination in Chile: averted cases, hospitalizations, ICU admissions, and deaths	Cases, hospitalizations, ICU admissions, deaths	Direct impact	<ol style="list-style-type: none"> <li>For each age group and each week, calculate the number of vaccinated individuals at risk by multiplying the number of individuals at risk with the proportion vaccinated.</li> <li>Calculate the incidence rate difference<sup>b</sup> between unvaccinated and vaccinated.</li> <li>Estimate the weekly averted outcomes by multiplying the number of vaccinated individuals at risk with the incidence rate difference.</li> <li>Sum the weekly averted outcomes across weeks and age groups.</li> </ol>

2023	Santos et al <sup>3</sup>	Brazil	Estimated COVID-19 severe cases and deaths averted in the first year of the vaccination campaign in Brazil: a retrospective observational study	Severe cases and deaths	Direct impact	Same as Haas et al.
2022	Kayano et al <sup>4</sup>	Japan	Number of averted COVID-19 cases and deaths attributable to reduced risk in vaccinated individuals in Japan	Cases and deaths	Direct impact	Similar to Brault et al, except that calculations in Steps 1 to 3 are stratified by age group, day, and sex.
2023	Kayano et al <sup>5</sup>	Japan (Tokyo)	Assessing the COVID-19 vaccination program during the Omicron variant (B.1.1.529) epidemic in early 2022, Tokyo	Infections	Direct impact, overall impact <sup>c</sup>	<b>Direct impact:</b> Similar to Brault et al <b>Overall impact:</b>
						1. Develop a transmission model based on a renewal process and calibrate the model to the cumulative incidence 2. Simulate the counterfactual trajectory under no vaccination by eliminating the vaccination from the fitted transmission model 3. Estimate the averted outcomes by taking the difference between observed outcomes and that under the counterfactual scenario of no vaccination
2021	Vilches et al <sup>6</sup>	USA	Estimating COVID-19 infections, hospitalizations, and deaths following the US vaccination campaigns during the pandemic	Infections, hospitalizations, and deaths	Overall impact	1. Calibrate an agent-based model to incidence data 2. Simulate the trajectory under scenarios with and without vaccination 3. Estimate the averted outcomes by taking the difference in outcomes between scenarios with and without vaccination
2022	Schneider et al <sup>7</sup>	USA	Impact of U.S. COVID-19 Vaccination Efforts: An Update on Averted Deaths, Hospitalizations, and Health	Infections, hospitalizations, and deaths	Overall impact	Similar to Vilches et al.

Care Costs Through March 2022						
2022	Gavish et al <sup>8</sup>	Israel	Population-level implications of the Israeli booster campaign to curtail COVID-19 resurgence	Cases, severe cases, and deaths	Overall impact <sup>d</sup>	Similar to Vilches et al., except that they considered scenarios with or without booster to investigate the outcomes averted by booster.
2021	Shoukat et al <sup>10</sup>	USA (New York City)	Lives saved and hospitalizations averted by COVID-19	Cases, hospitalizations, and deaths	Overall impact	<ol style="list-style-type: none"> <li>1. Calibrate an agent-based model to incidence data</li> <li>2. Simulate the trajectory under counterfactual scenario without vaccination</li> <li>3. Estimate the averted outcomes by taking the difference in outcomes between observed outcomes and that under the counterfactual scenario of no vaccination</li> </ol>
2022	Watson et al <sup>9</sup>	185 countries	Global impact of the first year of COVID-19 vaccination: a mathematical modelling study	Deaths <sup>e</sup>	Overall impact <sup>f</sup>	Similar to Vilches et al., except that they used an age-structured compartmental model

<sup>a</sup> Direct impact refers to the number of outcomes directly averted among vaccinated individuals; overall impact refers to the total number of outcomes averted among both vaccinated and unvaccinated individuals.

<sup>b</sup> Incidence rate among vaccinated individuals was implicitly conditional on not having developed the outcome until time of vaccination and thus was a hazard measure.

<sup>c</sup> The overall impact corresponds to what the authors referred to as the “total effect” in their paper, described as “the total effect at the population level, consisting of the vaccine-induced protection that is conferred directly and indirectly.”

<sup>d</sup> The authors also estimated the indirect effect of boosters.

<sup>e</sup> The authors also estimated all-cause mortality averted.

<sup>f</sup> The authors also estimated the number of outcomes averted via indirect effect.

## eAppendix 2. Defining individual, group, and population average potential outcomes

### 1 Mixed individual assignment strategy

Let  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$ , where  $\sum_{x=0}^{q+1} \rho_x = 1$ , be a parameterization that governs the distribution of vaccination times  $\mathbf{X}$ . Let  $K_x \equiv N \cdot \rho_x \equiv \sum_j I(X_j = x)$  for  $x \in \{0, \dots, q + 1\}$  and  $j \in \{1, \dots, N\}$ . Define  $\boldsymbol{\rho}$  to be a *mixed individual assignment strategy* if  $K_x$  is fixed under  $\boldsymbol{\rho}$ , with  $0 < K_x < N$  for all  $x \in \{0, \dots, q + 1\}$  and each of the  $\frac{N!}{\prod_{x=0}^{q+1} (K_x!)}$  possible individual assignments receiving equal probability.<sup>11,12</sup>

### 2 Type A parameterization with categorical treatment variable

Hudgens and Halloran<sup>11</sup> considered a mixed assignment strategy<sup>12</sup> for binary treatment variable (i.e., vaccination or non-vaccination) in defining the individual average potential outcome; their approach was referred to as Type A parameterization by VanderWeele and Tchetgen Tchetgen.<sup>13</sup> Here, we define the individual average potential outcomes using Type A parameterization for categorical treatment variable.

Assume  $\boldsymbol{\rho}$  is a mixed assignment strategy. Let  $\pi(\mathbf{X} = \mathbf{x}; \boldsymbol{\rho})$  denote the probability that the group is assigned with  $\mathbf{X}$  given parameter  $\boldsymbol{\rho}$ . The vaccination times  $\mathbf{X}$  are randomly assigned conditional on  $\{K_0, \dots, K_{q+1}\}$  with probability mass function:

$$\pi(\mathbf{X} = \mathbf{x}; \boldsymbol{\rho}) = \frac{\prod_{x=0}^{q+1} I(\sum_{j=1}^N I(X_j = x) = K_x)}{\frac{N!}{\prod_{x=0}^{q+1} (K_x!)}}$$

where  $\frac{N!}{\prod_{x=0}^{q+1} (K_x!)}$  is the number of ways to assign exactly  $K_0, \dots, K_{q+1}$  individuals to  $x = 0, \dots, q + 1$ , respectively.

### 3 Individual and group average potential outcomes

By the beginning of interval  $q + 1$ , the *individual average potential outcome*<sup>11</sup> for individual  $j$  with time of vaccination  $x$  in the group under strategy  $\boldsymbol{\rho}$  is:

$$\bar{Y}_{q+1,j}(x; \boldsymbol{\rho}) \equiv \sum_{\omega \in \mathcal{X}(N-1)} Y_{q+1,j}(\mathbf{x}_{-j} = \omega, x_j = x) \Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \omega | X_j = x)$$

where  $\Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \omega | X_j = x) = \frac{\pi(\mathbf{X}_{-j} = \omega, X_j = x; \boldsymbol{\rho})}{\sum_{\omega' \in \mathcal{X}(N-1)} \pi(\mathbf{X}_{-j} = \omega', X_j = x; \boldsymbol{\rho})}$ . See Definition 2 in VanderWeele and

Tchetgen Tchetgen for the analog with a binary treatment variable.<sup>13</sup> As they noted,<sup>13</sup> if  $\bar{Y}_{q+1,j}$  is defined under Type A parameterization, the proportion of other individuals assigned to, for example,  $x = 0$ , varies depending on whether individual  $j$  is assigned to  $x = 0$ . In other words, the proportion  $\rho_0$  for others is not held fixed in  $\bar{Y}_{q+1,j}(0; \boldsymbol{\rho})$  compared to  $\bar{Y}_{q+1,j}(k; \boldsymbol{\rho})$  for some  $k > 0$ , so that the so-called direct effect (which can be written as  $DE_{q+1}((k, 0); \boldsymbol{\rho}) = \bar{Y}_{q+1}(k; \boldsymbol{\rho}) - \bar{Y}_{q+1}(0; \boldsymbol{\rho}) = \frac{\sum_{j=1}^N (\bar{Y}_{q+1,j}(k; \boldsymbol{\rho}) - \bar{Y}_{q+1,j}(0; \boldsymbol{\rho}))}{N}$ )<sup>11</sup> does not merit the label “direct effect.” However, this issue is negligible under a large group size  $N$ .

The group average potential outcome is:

$$\bar{Y}_{q+1}(x; \boldsymbol{\rho}) = \frac{\sum_{j=1}^N \bar{Y}_{q+1,j}(x_j = x; \boldsymbol{\rho})}{N}.$$

### eAppendix 3. Causal estimand for outcomes directly averted by vaccination

#### 1 Extending the causal estimand to an arbitrary number of vaccination times

For  $q \in \mathbb{N}$ , the causal estimand for outcomes directly averted by  $\boldsymbol{\rho} = \{q+1, d; \rho_0, \dots, \rho_q\}$  where

$\sum_{x=0}^{q+1} \rho_x = 1$ , compared to no vaccination  $\boldsymbol{\Phi} = \{q+1, d; \mathbf{0}\}$  is:

*Definition S1 (Causal estimand for directly averted outcomes)*

$$\delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) = N \cdot \sum_{k=0}^q \rho_k \cdot [\bar{Y}_{q+1}(q+1; \boldsymbol{\rho}) - \bar{Y}_{q+1}(k; \boldsymbol{\rho})] \quad (\text{S1})$$

$$= N \cdot \sum_{k=0}^q \rho_k \cdot DE_{q+1}((q+1, k); \boldsymbol{\rho})$$

$\delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$  in equation (S1) can be re-written as:

$$\begin{aligned} & \delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) \\ &= N \cdot \sum_{k=0}^q \rho_k \cdot \left[ \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right] \\ &= N \cdot \left[ \rho_0 \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(0; \boldsymbol{\rho}) \right) + \rho_1 \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(1; \boldsymbol{\rho}) \right) + \rho_2 \right. \\ & \quad \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(2; \boldsymbol{\rho}) \right) + \dots + \rho_h \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(h; \boldsymbol{\rho}) \right) + \dots + \rho_q \\ & \quad \cdot \left. \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q; \boldsymbol{\rho}) \right) \right] \\ &= N \cdot \left[ \rho_0 \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(0; \boldsymbol{\rho}) \right) + \rho_1 \cdot \left( \sum_{l=2}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=2}^{q+1} \Delta \bar{Y}_l(1; \boldsymbol{\rho}) \right) + \rho_2 \right. \\ & \quad \cdot \left( \sum_{l=3}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=3}^{q+1} \Delta \bar{Y}_l(2; \boldsymbol{\rho}) \right) + \dots + \rho_{h-1} \cdot \left( \sum_{l=h}^{q+1} \Delta \bar{Y}_l(q+1; \boldsymbol{\rho}) - \sum_{l=h}^{q+1} \Delta \bar{Y}_l(h-1; \boldsymbol{\rho}) \right) + \dots + \rho_q \\ & \quad \cdot \left. \left( \Delta \bar{Y}_{q+1}(q+1; \boldsymbol{\rho}) - \Delta \bar{Y}_{q+1}(q; \boldsymbol{\rho}) \right) \right] \end{aligned}$$

$$\begin{aligned}
&= N \cdot \left[ \rho_0 \cdot (\Delta\bar{Y}_1(q+1; \mathbf{p}) - \Delta\bar{Y}_1(0; \mathbf{p})) + (\rho_0 + \rho_1) \cdot \left( \Delta\bar{Y}_2(q+1; \mathbf{p}) - \frac{\rho_0 \Delta\bar{Y}_2(0; \mathbf{p}) + \rho_1 \Delta\bar{Y}_2(1; \mathbf{p})}{\rho_0 + \rho_1} \right) + (\rho_0 + \rho_1 + \rho_2) \right. \\
&\quad \cdot \left( \Delta\bar{Y}_3(q+1; \mathbf{p}) - \frac{\rho_0 \Delta\bar{Y}_3(0; \mathbf{p}) + \rho_1 \Delta\bar{Y}_3(1; \mathbf{p}) + \rho_2 \Delta\bar{Y}_3(2; \mathbf{p})}{\rho_0 + \rho_1 + \rho_2} \right) + \dots + \left( \sum_{k=0}^{h-1} \rho_k \right) \\
&\quad \cdot \left( \Delta\bar{Y}_h(q+1; \mathbf{p}) - \frac{\sum_{k=0}^{h-1} \rho_k \Delta\bar{Y}_h(k; \mathbf{p})}{\sum_{k=0}^{h-1} \rho_k} \right) + \dots + \left( \sum_{k=0}^q \rho_k \right) \cdot \left( \Delta\bar{Y}_{q+1}(q+1; \mathbf{p}) - \frac{\sum_{k=0}^q \rho_k \Delta\bar{Y}_{q+1}(k; \mathbf{p})}{\sum_{k=0}^q \rho_k} \right) \left. \right] \\
&= N \cdot \sum_{l=1}^{q+1} \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \left( \Delta\bar{Y}_l(q+1; \mathbf{p}) - \frac{\sum_{k=0}^{l-1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p})}{\sum_{k=0}^{l-1} \rho_k} \right) \right]. \quad (\text{S2})
\end{aligned}$$

We assume individuals with  $x = l < q + 1$  have had the same historical probability of developing the outcome as those with  $x = q + 1$ , up until  $l$ . Therefore, from the third to the forth line,  $\Delta\bar{Y}_1(q+1; \mathbf{p})$  cancels out  $\Delta\bar{Y}_1(1; \mathbf{p})$ , since  $x = 1$  is unvaccinated before the beginning of interval 1; similarly,  $\sum_{l=1}^2 \Delta\bar{Y}_l(q+1; \mathbf{p})$  cancels out  $\sum_{l=1}^2 \Delta\bar{Y}_l(2; \mathbf{p})$ , since  $x = 2$  is unvaccinated before the beginning of interval 2, and so on for other vaccination groups. In words, equation (S2) is the proportion ever-vaccinated (i.e.,  $\sum_{k=0}^{l-1} \rho_k$ ) multiplied with the difference between the period incidence  $\Delta\bar{Y}_l(\cdot)$  among individuals assigned to no vaccination and the average period incidence among individuals ever-vaccinated, summed across time intervals  $l \in (1, \dots, q+1)$ .

Furthermore, as we assume individuals with  $x = l < q + 1$  have had the same historical probability of developing outcome as those with  $x = q + 1$ , up until  $l$ , we have  $\Delta\bar{Y}_l(q+1; \mathbf{p}) = \frac{\sum_{k=l}^{q+1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p})}{\sum_{k=l}^{q+1} \rho_k}$ , such that equation (S2) can be written as:

$$\delta_{q+1}^D(\mathbf{\Phi}, \mathbf{p}) = N \cdot \sum_{l=1}^{q+1} \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \left( \frac{\sum_{k=l}^{q+1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p})}{\sum_{k=l}^{q+1} \rho_k} - \frac{\sum_{k=0}^{l-1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p})}{\sum_{k=0}^{l-1} \rho_k} \right) \right] \quad (\text{S3})$$

$$= N \cdot \left[ \sum_{l=1}^{q+1} \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p})}{\sum_{k=l}^{q+1} \rho_k} - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} \rho_k \Delta\bar{Y}_l(k; \mathbf{p}) \right] \quad (\text{S4})$$

Writing  $\Delta\bar{Y}_l(q+1; \boldsymbol{\rho}) = \frac{\sum_{k=l}^{q+1} \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k}$  also has the advantage of preserving data because they

are equivalent in an ideal trial while  $\frac{\sum_{k=l}^{q+1} \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k}$  can be estimated with a larger sample.

## 2 An example: vaccination at two time points

To be explicit, consider vaccination at two time points. Recall equation (3) that defines outcomes directly averted by  $\boldsymbol{\rho} = \{2, d; \rho_0, \rho_1\}$ , compared to no vaccination  $\boldsymbol{\Phi} = \{2, d; \mathbf{0}\}$ :

$$\begin{aligned}
& \delta_2^D(\boldsymbol{\Phi}, \boldsymbol{\rho}) \\
&= N \cdot [\rho_0 \cdot (\bar{Y}_2(2; \boldsymbol{\rho}) - \bar{Y}_2(0; \boldsymbol{\rho})) + \rho_1 \cdot (\bar{Y}_2(2; \boldsymbol{\rho}) - \bar{Y}_2(1; \boldsymbol{\rho}))] \\
&= N \cdot \{\rho_0 \cdot [(\Delta\bar{Y}_1(2; \boldsymbol{\rho}) + \Delta\bar{Y}_2(2; \boldsymbol{\rho})) - (\Delta\bar{Y}_1(0; \boldsymbol{\rho}) + \Delta\bar{Y}_2(0; \boldsymbol{\rho}))] + \rho_1 \cdot [(\Delta\bar{Y}_1(2; \boldsymbol{\rho}) + \Delta\bar{Y}_2(2; \boldsymbol{\rho})) - (\Delta\bar{Y}_1(1; \boldsymbol{\rho}) + \Delta\bar{Y}_2(1; \boldsymbol{\rho}))]\} \\
&= N \cdot \{\rho_0 \cdot [(\Delta\bar{Y}_1(2; \boldsymbol{\rho}) + \Delta\bar{Y}_2(2; \boldsymbol{\rho})) - (\Delta\bar{Y}_1(0; \boldsymbol{\rho}) + \Delta\bar{Y}_2(0; \boldsymbol{\rho}))] + \rho_1 \cdot (\Delta\bar{Y}_2(2; \boldsymbol{\rho}) - \Delta\bar{Y}_2(1; \boldsymbol{\rho}))\} \\
&= N \cdot \left\{ \rho_0 \cdot (\Delta\bar{Y}_1(2; \boldsymbol{\rho}) - \Delta\bar{Y}_1(0; \boldsymbol{\rho})) + (\rho_0 + \rho_1) \left[ \Delta\bar{Y}_2(2; \boldsymbol{\rho}) - \left( \frac{\rho_0 \Delta\bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta\bar{Y}_2(1; \boldsymbol{\rho})}{\rho_0 + \rho_1} \right) \right] \right\} \\
\text{Assume } \Delta\bar{Y}_1(2; \boldsymbol{\rho}) &= \frac{\rho_1 \Delta\bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta\bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} \rightarrow \\
&= N \cdot \left\{ \rho_0 \cdot \left( \frac{\rho_1 \Delta\bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta\bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta\bar{Y}_1(0; \boldsymbol{\rho}) \right) + [(\rho_0 + \rho_1) \cdot \Delta\bar{Y}_2(2; \boldsymbol{\rho}) - (\rho_0 \Delta\bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta\bar{Y}_2(1; \boldsymbol{\rho}))] \right\} \\
&= N \cdot \sum_{l=1}^2 \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \left( \frac{\sum_{k=l}^2 \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^2 \rho_k} - \frac{\sum_{k=0}^{l-1} \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k} \right) \right] \\
&= N \cdot \left[ \sum_{l=1}^2 \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \left( \frac{\sum_{k=l}^2 \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^2 \rho_k} \right) - \sum_{l=1}^2 \sum_{k=0}^{l-1} \rho_k \Delta\bar{Y}_l(k; \boldsymbol{\rho}) \right]
\end{aligned}$$

## eAppendix 4 Conditions under which directly averted outcomes is a lower bound on overall averted outcomes

To identify conditions under which the number of directly averted outcomes among vaccinated individuals is a lower bound on the total number of outcomes averted, we first define the “overall impact” estimand to quantify outcomes averted in both vaccinated and unvaccinated individuals. Previously, we defined the overall impact estimand for vaccination at a single time point,<sup>14</sup> for which we now extend to multiple time points. First, we need to define the indirect, total, and overall *effects* for vaccination at multiple time points.

### 1 Indirect, total, and overall effects for vaccination at multiple time points

When vaccination occurs at multiple time points, there could be separate versions of indirect and total effects under each possible value  $x \in \{0, \dots, q + 1\}$ .

#### 1.1 Indirect effect

In general, indirect effect can be a contrast between two strategies  $\mathbf{p}$  and  $\tilde{\mathbf{p}}$ , conditional on  $x$ . In notation, the indirect effect comparing probability of having developed the outcome by beginning of interval  $q + 1$  for an individual assigned to vaccination time  $x$  when the group follows strategy  $\mathbf{p}$  versus  $\tilde{\mathbf{p}}$  is:

$$IE_{q+1}(x; (\mathbf{p}, \tilde{\mathbf{p}})) = \bar{Y}_{q+1}(x; \mathbf{p}) - \bar{Y}_{q+1}(x; \tilde{\mathbf{p}}). \quad (S5)$$

#### 1.2 Total effect

Total effect can be any contrast between any combinations of  $(x, \mathbf{p})$  and  $(x', \tilde{\mathbf{p}})$ . In notation, the total effect comparing probability of having developed the outcome by beginning of  $q + 1$  for an individual assigned to vaccination time  $x$  when the group follows strategy  $\mathbf{p}$  versus an individual with  $x'$  when the group follows strategy  $\tilde{\mathbf{p}}$  is:

$$TE_{q+1}((x, x'); (\mathbf{p}, \tilde{\mathbf{p}})) = \bar{Y}_{q+1}(x; \mathbf{p}) - \bar{Y}_{q+1}(x'; \tilde{\mathbf{p}}). \quad (S6)$$

By definition, total effect is the sum of direct and indirect effects:

$$\begin{aligned}
TE_{q+1}((x, x'); (\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}})) &= \bar{Y}_{q+1}(x; \boldsymbol{\rho}) - \bar{Y}_{q+1}(x'; \tilde{\boldsymbol{\rho}}) \\
&= \bar{Y}_{q+1}(x; \boldsymbol{\rho}) - \bar{Y}_{q+1}(x'; \boldsymbol{\rho}) + \bar{Y}_{q+1}(x'; \boldsymbol{\rho}) - \bar{Y}_{q+1}(x'; \tilde{\boldsymbol{\rho}}) \\
&= DE_{q+1}((x, x'); \boldsymbol{\rho}) + IE_{q+1}(x'; (\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}})). \quad (S7)
\end{aligned}$$

Hudgens and Halloran<sup>11</sup> also partitioned the total effect into the sum of direct and indirect effects for vaccination at a single time point.

### 1.3 Overall effect

The overall effect comparing probability of having developed the outcome by beginning of  $q + 1$  for a typical individual in the group following strategy  $\boldsymbol{\rho}$  versus  $\tilde{\boldsymbol{\rho}}$  is:

$$OE_{q+1}(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}) = \bar{Y}_{q+1}(\boldsymbol{\rho}) - \bar{Y}_{q+1}(\tilde{\boldsymbol{\rho}}). \quad (S8)$$

### 2 Partitioning overall effect when comparing no vaccination to some vaccination

Consider the comparison between no vaccination and a vaccination strategy  $\boldsymbol{\rho}$ , as our goal is to quantify the overall impact of the strategy. Previously, Sobel<sup>12</sup> and Hudgens and Halloran<sup>11</sup> noted that when comparing no vaccination to vaccination at baseline, overall effect is the weighted sum of total and indirect effects. Here, we extend the partitioning of overall effect to vaccination at multiple time points (Theorem S1).

*Theorem S1 (Overall effect partitioning for no vaccination vs.  $\boldsymbol{\rho}$ )* Let  $\boldsymbol{\Phi} = \{q + 1, d; \mathbf{0}\}$  for  $q \in \mathbb{N}$  and  $d \in \mathbb{Z}^+$ . Then

$$OE_{q+1}(\boldsymbol{\Phi}, \boldsymbol{\rho}) = \left[ \sum_{k=0}^q \rho_k \cdot TE_{q+1}((q + 1, k); (\boldsymbol{\Phi}, \boldsymbol{\rho})) \right] + \rho_{q+1} \cdot IE_{q+1}(q + 1; (\boldsymbol{\Phi}, \boldsymbol{\rho})).$$

*Proof.* For  $q \in \mathbb{N}$ , we have

$$OE_{q+1}(\boldsymbol{\Phi}, \boldsymbol{\rho}) = \bar{Y}_{q+1}(\boldsymbol{\Phi}) - \bar{Y}_{q+1}(\boldsymbol{\rho})$$

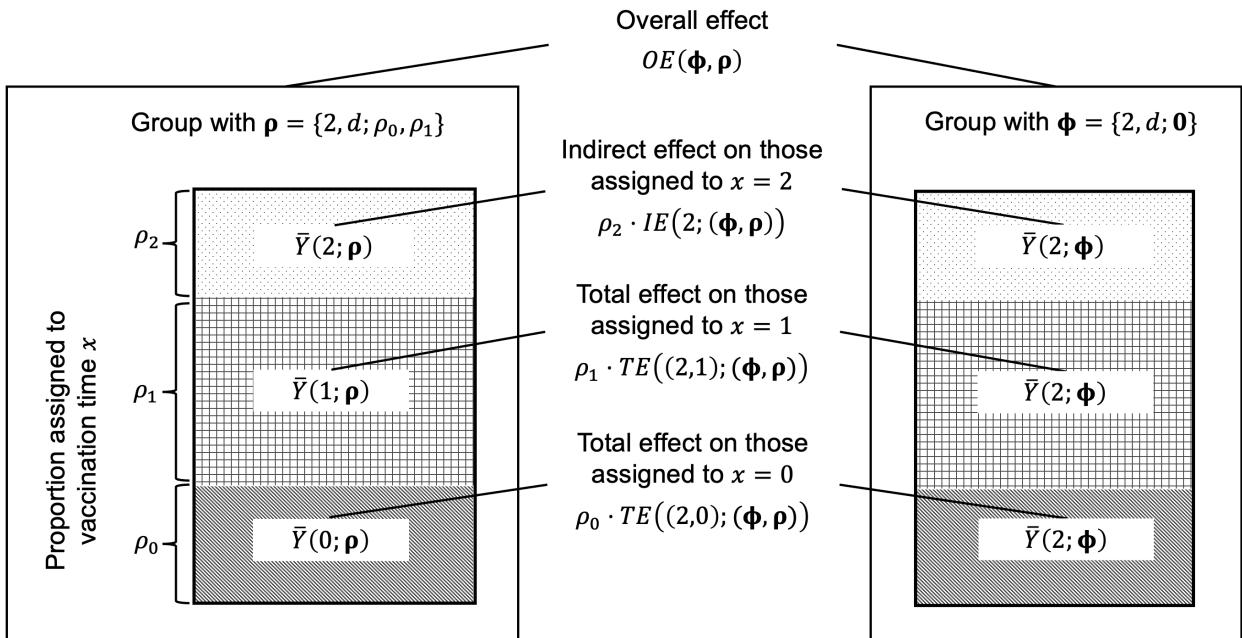
All is unvaccinated under  $\boldsymbol{\Phi} \rightarrow$

$$= \bar{Y}_{q+1}(q+1; \boldsymbol{\Phi}) - \bar{Y}_{q+1}(\boldsymbol{\rho})$$

By definition,  $\sum_{k=0}^{q+1} \rho_k = 1 \rightarrow$

$$\begin{aligned}
&= \left[ \sum_{k=0}^q \rho_k \cdot \bar{Y}_{q+1}(q+1; \boldsymbol{\Phi}) + \rho_{q+1} \cdot \bar{Y}_{q+1}(q+1; \boldsymbol{\Phi}) \right] \\
&\quad - \left[ \sum_{k=0}^q \rho_k \cdot \bar{Y}_{q+1}(k; \boldsymbol{\rho}) + \rho_{q+1} \cdot \bar{Y}_{q+1}(q+1; \boldsymbol{\rho}) \right] \\
&= \sum_{k=0}^q \rho_k \cdot (\bar{Y}_{q+1}(q+1; \boldsymbol{\Phi}) - \bar{Y}_{q+1}(k; \boldsymbol{\rho})) + \rho_{q+1} \cdot (\bar{Y}_{q+1}(q+1; \boldsymbol{\Phi}) - \bar{Y}_{q+1}(q+1; \boldsymbol{\rho})) \\
&= \sum_{k=0}^q \rho_k \cdot TE_{q+1}((q+1, k); (\boldsymbol{\Phi}, \boldsymbol{\rho})) + \rho_{q+1} \cdot IE_{q+1}(q+1; (\boldsymbol{\Phi}, \boldsymbol{\rho}))
\end{aligned}$$

Consider  $q+1 = 2$ . We have  $OE_2(\boldsymbol{\Phi}, \boldsymbol{\rho}) = \rho_0 TE_2((2,0); (\boldsymbol{\Phi}, \boldsymbol{\rho})) + \rho_1 TE_2((2,1); (\boldsymbol{\Phi}, \boldsymbol{\rho})) + \rho_2 IE_2(2; (\boldsymbol{\Phi}, \boldsymbol{\rho}))$ . The partitioning is graphically illustrated in **Figure S1**.



**FIGURE S1 | Graphical illustration on partitioning overall effect.** The two rectangles represent a pair of counterfactuals wherein the group follow strategy  $\boldsymbol{\rho} = \{2, d; \rho_0, \rho_1\}$  or  $\boldsymbol{\Phi} = \{2, d; \mathbf{0}\}$ . Individuals fall into three categories based on their assigned vaccination time: 1) The dotted region represents those who are assigned to  $x = 2$  and for whom  $IE(2; \boldsymbol{\Phi}, \boldsymbol{\rho})$  is the difference between the counterfactuals; 2) the gridded region represents those who are assigned to  $x = 1$  and for whom  $TE((2,1); (\boldsymbol{\Phi}, \boldsymbol{\rho}))$  is the difference between the counterfactuals; and 3) the stripped region represents those who are assigned to  $x = 0$  and for whom  $TE((2,0); (\boldsymbol{\Phi}, \boldsymbol{\rho}))$  is the difference between the counterfactuals. Theorem S1 shows that  $OE(\boldsymbol{\Phi}, \boldsymbol{\rho})$  is a weighted average of three effects: 1)  $IE(2; (\boldsymbol{\Phi}, \boldsymbol{\rho}))$ , 2)  $TE((2,1); (\boldsymbol{\Phi}, \boldsymbol{\rho}))$ , and 3)  $TE((2,0); (\boldsymbol{\Phi}, \boldsymbol{\rho}))$ , each weighted by the proportion of individuals for whom the effect is in operation respectively: 1)  $\rho_2$  for those assigned to  $x = 2$ , 2)  $\rho_1$  for those assigned to  $x = 1$ , and 3)  $\rho_0$  for those assigned to  $x = 0$ . The time interval notation suppressed to reduce clutter.

### 3 Defining and partitioning overall impact for vaccination at multiple time points

#### 3.1 Overall impact

Now, we use the above definition of overall effect to define the overall impact estimand and apply Theorem S1 to partition overall impact into components of direct and indirect effects.

The total number of outcomes averted among both vaccinated and unvaccinated individuals by strategy  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$ , compared to no vaccination  $\boldsymbol{\Phi} = \{q + 1, d; \mathbf{0}\}$ , is:

*Definition S2 (Overall impact estimand for quantifying the total number of outcomes averted)*

$$\delta_{q+1}^O(\boldsymbol{\Phi}, \boldsymbol{\rho}) = N \cdot OE_{q+1}(\boldsymbol{\Phi}, \boldsymbol{\rho}).$$

$\delta_{q+1}^O(\boldsymbol{\Phi}, \boldsymbol{\rho})$  addresses the causal question: how many outcomes have been averted among both vaccinated and unvaccinated individuals under strategy  $\boldsymbol{\rho}$  compared to no vaccination  $\boldsymbol{\Phi}$ ? The mathematical modeling studies<sup>5–10</sup> listed in **Table S1** have targeted  $\delta_{q+1}^O(\boldsymbol{\Phi}, \boldsymbol{\rho})$  to estimate total

outcomes averted by simulating the epidemic trajectory under no vaccination ( $\Phi$ ) and comparing it with the trajectory under a particular strategy  $\rho$  (or with the observed outcomes).

### 3.2 Overall impact partitioning

By Theorem S1 and the definition from equation (S7)  $TE_{q+1}((q+1, k); (\Phi, \rho)) = IE_{q+1}(q +$

$1; (\Phi, \rho)) + DE_{q+1}((q+1, k); \rho)$  for  $k \in \{0, \dots, q\}$ , we decompose  $\delta_{q+1}^O(\Phi, \rho)$ :

$$\delta_{q+1}^O(\Phi, \rho) = N \cdot OE_{q+1}(\Phi, \rho)$$

$$\begin{aligned} &= N \left[ \sum_{k=0}^q \rho_k \cdot TE_{q+1}((q+1, k); (\Phi, \rho)) + \rho_{q+1} \cdot IE_{q+1}(q+1; (\Phi, \rho)) \right] \\ &= N \left[ \sum_{k=0}^q \rho_k \cdot \left( DE_{q+1}((q+1, k); \rho) + IE_{q+1}(q+1; (\Phi, \rho)) \right) + \rho_{q+1} \cdot IE_{q+1}(q+1; (\Phi, \rho)) \right] \\ &= N \left[ \left( \sum_{k=0}^q \rho_k \cdot DE_{q+1}((q+1, k); \rho) \right) + IE_{q+1}(q+1; (\Phi, \rho)) \right] \quad (\text{S9}) \end{aligned}$$

Substituting the definition  $\delta_{q+1}^D(\Phi, \rho) = N \cdot \sum_{k=0}^q \rho_k \cdot DE_{q+1}((q+1, k); \rho)$  into the last line of equation (S9), we have:

$$\delta_{q+1}^O(\Phi, \rho) = \delta_{q+1}^D(\Phi, \rho) + N \cdot IE_{q+1}(q+1; (\Phi, \rho))$$

Therefore,  $\delta_{q+1}^O(\Phi, \rho) \geq \delta_{q+1}^D(\Phi, \rho)$  if and only if  $IE_{q+1}(q+1; (\Phi, \rho)) \geq 0$ .

## eAppendix 5. Unbiased estimator for the causal estimand in a one-stage randomized controlled trial

Recall equation (S1):  $\delta_{q+1}^D(\Phi, \rho) = N \cdot \sum_{k=0}^q \rho_k \cdot [\bar{Y}_{q+1}(q+1; \rho) - \bar{Y}_{q+1}(k; \rho)]$ . To identify  $\delta_{q+1}^D(\Phi, \rho)$ , consider the following assumptions.

*Assumption 1.*  $\rho$  is a mixed assignment strategy.

Recall the definition of mixed assignment strategy in **eAppendix 2**:  $\rho$  is a *mixed individual assignment strategy* if  $K_x$  is fixed under  $\rho$ , with  $0 < K_x < N$  for all  $x \in \{0, \dots, q+1\}$  and each of the  $\frac{N!}{\prod_{x=0}^{q+1} (K_x!)}$  possible individual assignments receiving equal probability.<sup>11,12</sup>

*Assumption 2 (consistency).*<sup>26</sup>

$$\text{If } \mathbf{X} = \mathbf{x}, \text{ then } Y_{q+1,j}(\mathbf{x}) = Y_{q+1,j}$$

for all  $j$ .

To identify  $\delta_{q+1}^D(\Phi, \rho)$ , we can make use of the following theorem.

*Theorem S2 (Unbiased estimator for the causal estimand in a one-stage RCT).* Let  $\rho = \{q+1, d; \rho_0, \dots, \rho_q\}$ , where  $\sum_{x=0}^{q+1} \rho_x = 1$ , and  $\Phi = \{q+1, d; \mathbf{0}\}$  for  $q \in \mathbb{N}$  and  $d \in \mathbb{Z}^+$ . Let

$$\widehat{\delta}_{q+1}^D(\Phi, \rho) = N \cdot \sum_{k=0}^q \rho_k \cdot [\hat{Y}_{q+1}(q+1; \rho) - \hat{Y}_{q+1}(k; \rho)] \quad (\text{S10})$$

where  $\hat{Y}_{q+1}(x; \rho) = \frac{\sum_{j=1}^N Y_{q+1,j}(\mathbf{x}) I[X_j=x]}{\sum_{j=1}^N I[X_j=x]}$  for  $x \in \{0, \dots, q+1\}$ . That is,  $\hat{Y}_{q+1}(x; \rho)$  is the average of

observed outcomes for individuals assigned with  $x$  under strategy  $\rho$ . Under Assumpstions 1 and 2,

$$E \left[ \widehat{\delta}_{q+1}^D(\Phi, \rho) \right] = \delta_{q+1}^D(\Phi, \rho).$$

*Proof:* First, we expand LHS:

$$E\left[\widehat{\delta_{q+1}}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})\right] = E\left[N \cdot \sum_{k=0}^q \rho_k \cdot [\widehat{Y}_{q+1}(q+1; \boldsymbol{\rho}) - \widehat{Y}_{q+1}(k; \boldsymbol{\rho})]\right]$$

Linearity of expectation  $\rightarrow$

$$= N \cdot \sum_{k=0}^q [\rho_k \cdot (E[\widehat{Y}_{q+1}(q+1; \boldsymbol{\rho})] - E[\widehat{Y}_{q+1}(k; \boldsymbol{\rho})])] \quad (\text{S11}).$$

We largely repeat the proof in A.1 of Hudgens and Halloran<sup>11</sup> to evaluate  $E[\widehat{Y}_{q+1}(x; \boldsymbol{\rho})]$  for  $x \in \{0, \dots, q+1\}$ . Without loss of generality, let  $x = 0$ .

$$E[\widehat{Y}_{q+1}(0; \boldsymbol{\rho})] = E\left[\frac{\sum_{j=1}^N Y_{q+1,j}(\mathbf{X}) I[X_j = 0]}{\sum_{j=1}^N I[X_j = 0]}\right]$$

Under Assumption 1,  $K_0 \equiv N \cdot \rho_0$  is fixed by design, and under Assumption 2 (Consistency), we have:

$$E[\widehat{Y}_{q+1}(0; \boldsymbol{\rho})] = \frac{1}{K_0} \sum_{j=1}^N \sum_{\mathbf{s} \in \mathcal{X}(N)} \Pr_{\boldsymbol{\rho}}(\mathbf{X} = \mathbf{s}) Y_{q+1,j}(\mathbf{s}) I[x_j = 0].$$

Any  $\mathbf{s}$  such that  $x_j \neq 0$  does not contribute to the summation, so that we can write:

$$\begin{aligned} & E[\widehat{Y}_{q+1}(0; \boldsymbol{\rho})] \\ &= \frac{1}{K_0} \sum_{j=1}^N \sum_{\boldsymbol{\omega} \in \mathcal{X}(N-1)} \Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \boldsymbol{\omega}, X_j = 0) Y_{q+1,j}(\mathbf{x}_{-j} = \boldsymbol{\omega}, x_j = 0) \\ &= \frac{1}{K_0} \sum_{j=1}^N \sum_{\boldsymbol{\omega} \in \mathcal{X}(N-1)} \Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \boldsymbol{\omega} | X_j = 0) \Pr_{\boldsymbol{\rho}}(X_j = 0) Y_{q+1,j}(\mathbf{x}_{-j} = \boldsymbol{\omega}, x_j = 0). \end{aligned}$$

Under Assumption 1,  $\Pr_{\boldsymbol{\rho}}(X_j = 0) = \rho_0 = K_0/N$ , implying

$$E[\widehat{Y}_{q+1}(0; \boldsymbol{\rho})] = \frac{1}{N} \sum_{j=1}^N \sum_{\boldsymbol{\omega} \in \mathcal{X}(N-1)} \Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \boldsymbol{\omega} | X_j = 0) Y_{q+1,j}(\mathbf{x}_{-j} = \boldsymbol{\omega}, x_j = 0)$$

Recall  $\bar{Y}_{q+1,j}(0; \boldsymbol{\rho}) \equiv \sum_{\boldsymbol{\omega} \in \mathcal{X}(N-1)} \Pr_{\boldsymbol{\rho}}(\mathbf{X}_{-j} = \boldsymbol{\omega} | X_j = x) Y_{q+1,j}(\mathbf{x}_{-j} = \boldsymbol{\omega}, x_j = 0) \rightarrow$

$$= \frac{1}{N} \sum_{j=1}^N \bar{Y}_{q+1,j}(0; \boldsymbol{\rho})$$

$$\text{Recall } \bar{Y}_{q+1}(0; \boldsymbol{\rho}) = \frac{\sum_{j=1}^N \bar{Y}_{q+1,j}(0; \boldsymbol{\rho})}{N} \rightarrow$$

$$= \bar{Y}_{q+1}(0; \boldsymbol{\rho}).$$

Recall equation (S11):

$$E \left[ \widehat{\delta_{q+1}^D}(\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = N \cdot \sum_{k=0}^q [\rho_k \cdot (E[\hat{Y}_{q+1}(q+1; \boldsymbol{\rho})] - E[\hat{Y}_{q+1}(k; \boldsymbol{\rho})])]$$

Substituting from above  $\rightarrow$

$$= N \cdot \sum_{k=0}^q \rho_k [\bar{Y}_{q+1}(q+1; \boldsymbol{\rho}) - \bar{Y}_{q+1}(k; \boldsymbol{\rho})]$$

Recall equation (S1)  $\rightarrow$

$$= \delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$$

## eAppendix 6. Extending the hazard difference estimator for vaccination at more than two time points

### 1 Hazard difference estimator for quantifying directly averted outcomes

The hazard difference estimator for strategy  $\boldsymbol{\rho} = \{q+1, d; 0, \dots, q\}$  for  $q \in \mathbb{N}$  can be defined as follows:

$$\widehat{\delta}_{q+1}^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho}) = \sum_{l=1}^{q+1} [\widehat{N}_{l-1}^v(\boldsymbol{\rho}) \cdot (\widehat{h}_l^u(\boldsymbol{\rho}) - \widehat{h}_l^v(\boldsymbol{\rho}))]$$

where  $\widehat{N}_{l-1}^v(\boldsymbol{\rho}) = N \sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))$ ,  $\widehat{h}_l^u(\boldsymbol{\rho}) = \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}$ , and  $\widehat{h}_l^v(\boldsymbol{\rho}) = \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}$ .

### 2 Bias of hazard difference estimator relative to the causal estimand

#### 2.1 Analytical comparison

##### 2.1.1 Bias of hazard difference estimator for an arbitrary number of vaccination times

To examine the bias of  $\widehat{\delta}_{q+1}^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho})$  relative to  $\delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$ , we first expand  $E[\widehat{\delta}_{q+1}^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho})]$  as follows:

$$\begin{aligned} E[\widehat{\delta}_{q+1}^{D^*}(\boldsymbol{\Phi}, \boldsymbol{\rho})] &= E\left[\sum_{l=1}^{q+1} [\widehat{N}_{l-1}^v(\boldsymbol{\rho}) \cdot (\widehat{h}_l^u(\boldsymbol{\rho}) - \widehat{h}_l^v(\boldsymbol{\rho}))]\right] \\ &= E\left[\sum_{l=1}^{q+1} \left[N \cdot \sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho})) \cdot \left(\frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))} - \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}\right)\right]\right] \\ &= N \left\{ E\left[\sum_{l=1}^{q+1} \left[\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho})) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))} - \sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})\right]\right]\right\} \end{aligned}$$

Linearity of expectation  $\rightarrow$

$$= N \left\{ \sum_{l=1}^{q+1} E\left[\left[\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))\right] \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}\right] - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} E[\rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})] \right\}$$

By a proof parallel with that of Theorem S2, we have:

$$E \left[ \widehat{\delta_{q+1}^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = N \left\{ \sum_{l=1}^{q+1} E \left[ \left( \sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho})) \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))} \right] - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right\}$$

Rearranging terms  $\rightarrow$

$$= N \left\{ \sum_{l=1}^{q+1} E \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \frac{\frac{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k} \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k} \right] - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right\} \quad (S12)$$

If  $\frac{\frac{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k}}{\frac{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=l}^{q+1} \rho_k}} = 1$  (i.e., vaccination has no effect), then we can write equation (S12) as:

$$E \left[ \widehat{\delta_{q+1}^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = N \left\{ \sum_{l=1}^{q+1} E \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k} \right] - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right\}$$

By a proof parallel with that of Theorem S2, we have:

$$E \left[ \widehat{\delta_{q+1}^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = N \left\{ \sum_{l=1}^{q+1} \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k} - \sum_{l=1}^{q+1} \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right\}$$

Recall equation[S4]  $\rightarrow$

$$= \delta_{q+1}^D (\boldsymbol{\Phi}, \boldsymbol{\rho}).$$

However, if  $\frac{\frac{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k}}{\frac{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=l}^{q+1} \rho_k}} \neq 1$ , then  $E \left[ \widehat{\delta_{q+1}^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] \neq \delta_{q+1}^D (\boldsymbol{\Phi}, \boldsymbol{\rho})$ . That is,

$\widehat{\delta_{q+1}^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho})$  is a biased estimator if the survival among vaccinated individuals is different from that among unvaccinated individuals. In an ideal RCT, the differential survival is due to a non-null effect of vaccination.

### 2.1.2 An example: vaccination at two time points

From equation (S12), we have:

$$E \left[ \widehat{\delta_2^D}^* (\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = E \left[ \widehat{N}_0^v(\boldsymbol{\rho}) \left( \widehat{h}_1^u(\boldsymbol{\rho}) - \widehat{h}_1^v(\boldsymbol{\rho}) \right) + \widehat{N}_1^v(\boldsymbol{\rho}) \left( \widehat{h}_2^u(\boldsymbol{\rho}) - \widehat{h}_2^v(\boldsymbol{\rho}) \right) \right]$$

$$\begin{aligned}
&= N \left\{ \sum_{l=1}^2 E \left[ \left( \sum_{k=0}^{l-1} \rho_k \right) \cdot \frac{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=l}^{l-1} \rho_k} \cdot \frac{\sum_{k=l}^2 \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^2 \rho_k} \right] - \sum_{l=1}^2 \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right\} \\
&= N \left\{ E \left[ \rho_0 \cdot \frac{\rho_0 (1 - \hat{Y}_0(0; \boldsymbol{\rho}))}{\rho_0 + \rho_1} \cdot \frac{\rho_1 \Delta \hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \hat{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} \right] \right. \\
&\quad + E \left[ (\rho_0 + \rho_1) \cdot \frac{\frac{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))}{\rho_0 + \rho_1} \cdot \frac{\rho_2 \Delta \hat{Y}_2(2; \boldsymbol{\rho})}{\rho_2}}{\frac{\rho_2 (1 - \hat{Y}_0(2; \boldsymbol{\rho}))}{\rho_2}} \right. \\
&\quad \left. \left. - [\rho_0 \Delta \bar{Y}_1(0; \boldsymbol{\rho}) + (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho}))] \right] \right\}
\end{aligned}$$

since  $\hat{Y}_0 \equiv 0 \rightarrow$

$$\begin{aligned}
&= N \left\{ E \left[ \rho_0 \cdot \frac{\rho_1 \Delta \hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \hat{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} \right] + E \left[ (\rho_0 + \rho_1) \cdot \frac{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))}{(\rho_0 + \rho_1) (1 - \hat{Y}_1(2; \boldsymbol{\rho}))} \cdot \Delta \hat{Y}_2(2; \boldsymbol{\rho}) \right] \right. \\
&\quad \left. - [\rho_0 \Delta \bar{Y}_1(0; \boldsymbol{\rho}) + (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho}))] \right\}
\end{aligned}$$

By a proof parallel with that of Theorem S2, we have:

$$\begin{aligned}
&= N \left\{ \rho_0 \cdot \frac{\rho_1 \Delta \bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} + E \left[ (\rho_0 + \rho_1) \cdot \frac{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))}{(\rho_0 + \rho_1) (1 - \hat{Y}_1(2; \boldsymbol{\rho}))} \cdot \Delta \hat{Y}_2(2; \boldsymbol{\rho}) \right] \right. \\
&\quad \left. - [\rho_0 \Delta \bar{Y}_1(0; \boldsymbol{\rho}) + (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho}))] \right\}
\end{aligned}$$

rearranging terms  $\rightarrow$

$$\begin{aligned}
&= N \left\{ \rho_0 \cdot \left( \frac{\rho_1 \Delta \bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \bar{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta \bar{Y}_1(0; \boldsymbol{\rho}) \right) + E \left[ (\rho_0 + \rho_1) \cdot \frac{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))}{(\rho_0 + \rho_1) (1 - \hat{Y}_1(2; \boldsymbol{\rho}))} \cdot \Delta \hat{Y}_2(2; \boldsymbol{\rho}) \right] \right. \\
&\quad \left. - (\rho_0 \Delta \bar{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \bar{Y}_2(1; \boldsymbol{\rho})) \right\} \quad (\text{S13})
\end{aligned}$$

From the preceding section, we conclude that  $E[\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho})] \neq \delta_2^D(\boldsymbol{\phi}, \boldsymbol{\rho})$  if  $\frac{\rho_0(1-\hat{Y}_1(0;\boldsymbol{\rho})) + \rho_1(1-\hat{Y}_1(1;\boldsymbol{\rho}))}{(\rho_0+\rho_1)(1-\hat{Y}_1(2;\boldsymbol{\rho}))} \neq 1$  (i.e., vaccination has a non-null effect).

## 2.2 Simulations

Here we consider strategy  $\boldsymbol{\rho}'' = \{99,7; \mathbf{0.01}\}$ , where 1% of the individuals are randomized to vaccination at  $x \in \{0,1,\dots,98\}$ , or to no vaccination throughout (i.e.,  $x = 99$ ). To illustrate the bias of hazard difference estimator relative to the causal estimand, we simulate the epidemic under the same nine scenarios as described in **Table 1**. Compartments are stratified by  $x \in \{0,1,\dots,99\}$ , such that within each stratum  $x$ , we specify a continuous-time SIRD model in term of continuous time  $t$  as follows:

$$\left. \begin{aligned} \frac{dS_x(t)}{dt} &= -\theta_x(t) \cdot \lambda(t) \cdot S_x(t) \\ \frac{dI_x(t)}{dt} &= \theta_x(t) \cdot \lambda(t) \cdot S_x(t) - \gamma \cdot I_x(t) \\ \frac{dR_x(t)}{dt} &= (1 - \kappa_x(t) \cdot \mu) \cdot \gamma \cdot I_x(t) \\ \frac{dD_x(t)}{dt} &= \kappa_x(t) \cdot \mu \cdot \gamma \cdot I_x(t) \end{aligned} \right\} \quad (\text{S14})$$

where  $\lambda(t) = \beta \cdot \frac{\sum_{x=0}^{q+1} I_x(t)}{N(t)}$  and  $N(t) =$  the sum of all individuals alive at  $t$ . For simulation, the model was discretized to day time-steps from Day 0 (i.e., the beginning of interval 0) to Day 693 (i.e., the beginning of interval 99). All model parameters are the same as **Table S4**, except that we use time-varying parameter  $\theta_x(t)$  to parameterize  $1-\text{VE}_{\text{inf}} / 100\%$ :

$$\theta_x(t) = \begin{cases} 1 & \text{if } t < \text{beginning of interval } x \\ \theta & \text{if } t \geq \text{beginning of interval } x \end{cases}$$

for vaccination interval  $x$  and time  $t$ , where  $\theta_x(\cdot)$  can be 10% or 0%, depending on the scenario as outlined in **Table 1**. Similarly, we use time-varying parameter  $\kappa_x(t)$  to parameterize  $1-\text{VE}_{\text{death}} / 100\%$ :

$$\kappa_x(t) = \begin{cases} 1 & \text{if } t < \text{beginning of interval } x \\ \kappa & \text{if } t \geq \text{beginning of interval } x \end{cases}$$

for vaccination interval  $x$  and time  $t$ , where  $\kappa_x(\cdot)$  can be 10% or 0%, depending on the scenario as outlined in **Table 1**.

Initial values for the simulation are described in **Table S2**.

**TABLE S2** | Initial values of the simulations under strategy  $\rho'' = \{99,7; \mathbf{0.01}\}$ .

Variable	Initial condition(s)	Definition
$N(0)$	300,000	Number alive at baseline
$S_x(0)$	$(300,000 - 300) * 0.01$	Number of susceptible individuals assigned to receive vaccination at the beginning of interval $x$
$I_x(0)$	$300 * 0.01$	Number of infectious individuals assigned to receive vaccination at the beginning of interval $x$
$R_x(0)$	0	Number of recovered individuals assigned to receive vaccination at the beginning of interval $x$
$D_x(0)$	0	Number of individuals assigned to receive vaccination at the beginning of interval $x$ who died due to infection

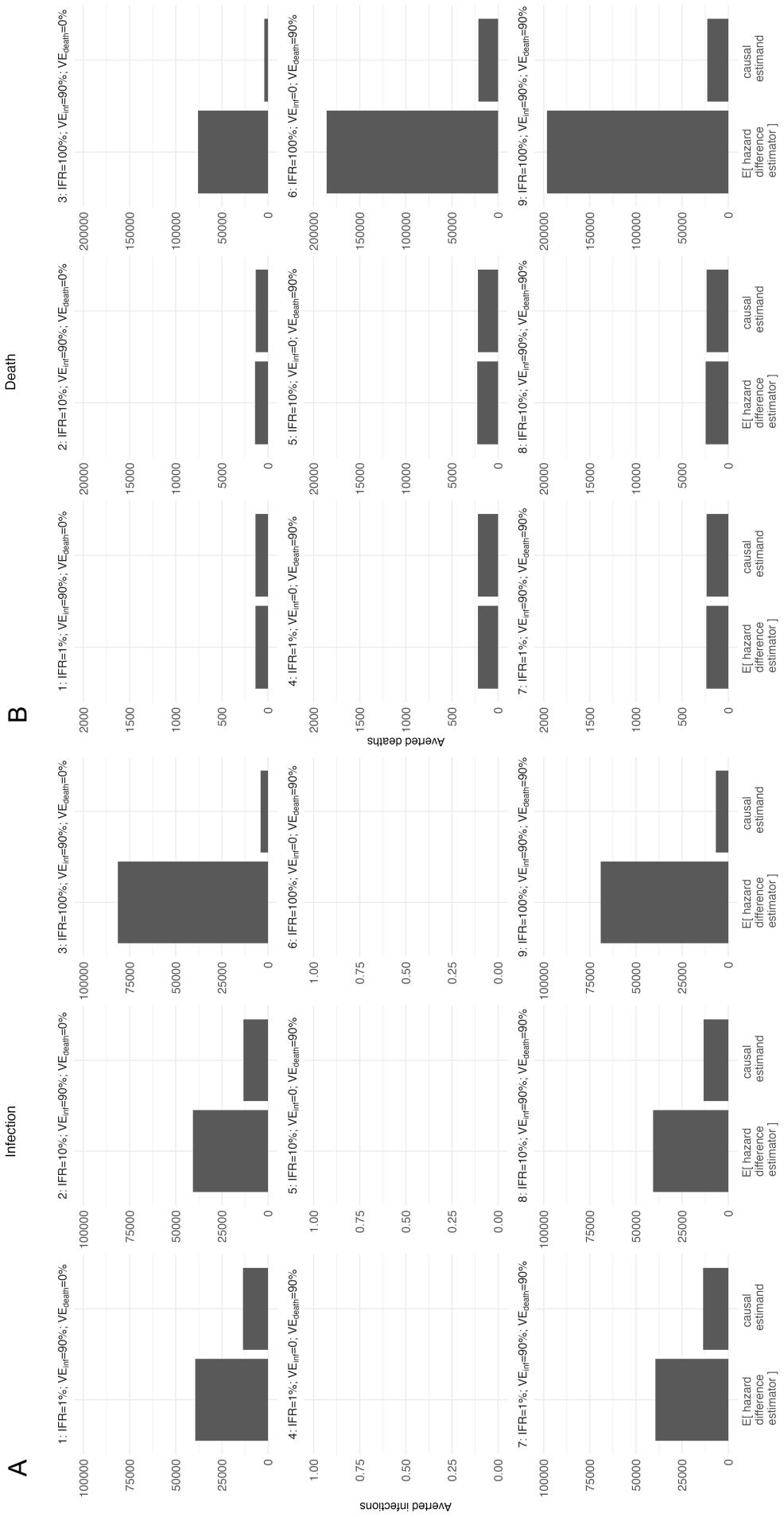
**Figure S2** compares the expected value of hazard difference estimator and the causal estimand for directly averted infections and deaths under strategy  $\rho'' = \{99,7; \mathbf{0.01}\}$  before the beginning of Day 693, while **Table S3** shows the absolute and percentage biases. Consistent with the derivation above, the hazard difference estimator overestimates averted outcomes when  $VE_{inf} > 0$ , due to the lower survival of not-yet-vaccinated individuals compared to ever-vaccinated

individuals (i.e.,  $\frac{\sum_{k=0}^{l-1} \rho''_k (1 - \bar{\gamma}_{l-1}(k; \rho''))}{\sum_{k=0}^{l-1} \rho''_k} > 1$  in equation [S12]).

**TABLE S3** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand (reference) for averted outcomes under strategy  $\rho''$  under scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ ), as outlined in **Table 1**.

	<b>Scenario 1</b>	<b>Scenario 2</b>	<b>Scenario 3</b>	<b>Scenario 4</b>	<b>Scenario 5</b>	<b>Scenario 6</b>	<b>Scenario 7</b>	<b>Scenario 8</b>	<b>Scenario 9</b>
	IFR: 1%, $VE_{inf} =$ 90%, $VE_{death} = 0\%$	IFR: 10%, $VE_{inf} =$ 90%, $VE_{death} = 0\%$	IFR: 100%, $VE_{inf} =$ 90%, $VE_{death} = 0\%$	IFR: 1%, $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	IFR: 100%, $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	IFR: 1%, $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	IFR: 10%, $VE_{inf} = 90\%$ , $VE_{death} = 90\%$	IFR: 100%, $VE_{inf} = 90\%$ , $VE_{death} = 90\%$	IFR: 100%, $VE_{inf} = 90\%$ , $VE_{death} = 90\%$
<b>Averted infections <sup>a</sup></b>	25930.18 (191.06%)	27362.86 (204.35%)	77260.67 (1883.18%)	0.00	0.00	0.00	25926.36 (191.03%)	27319.99 (203.97%)	62417.26 (931.8%)
<b>Averted deaths <sup>a</sup></b>	0.57 (0.42%)	60.81 (4.54%)	71795.51 (1749.97%)	0.76 (0.35%)	81.7 (3.74%)	164255.82 (763.85%)	0.83 (0.35%)	89 (3.77%)	173931.79 (769.03%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places.



**FIGURE S2 |** Infections (A) and deaths (B) directly averted by strategy  $\rho'' = \{99, 7, 0, 01\}$  under scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ ), as outlined in **Table 1**.

## eAppendix 7. An alternative estimator for quantifying directly averted outcomes using RCT data aggregated by vaccination status

Here, we show that, for averted outcomes, the causal estimand can be identified with an alternative unbiased estimator using data aggregated by vaccination status through an expression similar (but not identical) to that of the hazard difference estimator. While this estimator may appear unnecessary for RCTs, where data are disaggregated by vaccination time and equation (S10) can be directly applied, it has important implications for empirical analyses that rely on aggregated data from national vaccine systems, as discussed further in **eAppendix 13**.

*Theorem S3 (Unbiasedness of an alternative estimator for averted outcomes).* Let  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$ , where  $\sum_{x=0}^{q+1} \rho_x = 1$ , and  $\boldsymbol{\Phi} = \{q + 1, d; \mathbf{0}\}$  for  $q \in \mathbb{N}$  and  $d \in \mathbb{Z}^+$ . Let

$$\widehat{\delta}_{q+1}^{D'}(\boldsymbol{\Phi}, \boldsymbol{\rho}) = \sum_{l=1}^{q+1} \left( \frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \widehat{N}_{l-1}^u(\boldsymbol{\rho}) \widehat{h}_l^u(\boldsymbol{\rho}) - \widehat{N}_{l-1}^v(\boldsymbol{\rho}) \widehat{h}_l^v(\boldsymbol{\rho}) \right). \quad (\text{S15})$$

where  $\widehat{N}_{l-1}^u(\boldsymbol{\rho}) = N \sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))$ ,  $\widehat{N}_{l-1}^v(\boldsymbol{\rho}) = N \sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))$ ,  $\widehat{h}_l^u(\boldsymbol{\rho}) =$

$$\frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}, \text{ and } \widehat{h}_l^v(\boldsymbol{\rho}) = \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))}.$$

Under Assumptions 1 and 2 from eAppendix 5,  $E \left[ \widehat{\delta}_{q+1}^{D'}(\boldsymbol{\Phi}, \boldsymbol{\rho}) \right] = \delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$ .

*Proof:*

$$E[\widehat{\delta}_{q+1}^{D'}(\boldsymbol{\Phi}, \boldsymbol{\rho})] = E \left[ \sum_{l=1}^{q+1} \left( \frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \widehat{N}_{l-1}^u(\boldsymbol{\rho}) \widehat{h}_l^u(\boldsymbol{\rho}) - \widehat{N}_{l-1}^v(\boldsymbol{\rho}) \widehat{h}_l^v(\boldsymbol{\rho}) \right) \right]$$

Substituting  $\widehat{N}_{l-1}^u(\boldsymbol{\rho})$ ,  $\widehat{h}_l^u(\boldsymbol{\rho})$ ,  $\widehat{N}_{l-1}^v(\boldsymbol{\rho})$ , and  $\widehat{h}_l^v(\boldsymbol{\rho})$  from above  $\rightarrow$

$$\begin{aligned} &= \sum_{l=1}^{q+1} E \left[ \left( N \cdot \frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \cdot \sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho})) \cdot \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))} - N \cdot \sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho})) \right. \right. \\ &\quad \left. \left. \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \boldsymbol{\rho}))} \right) \right] \end{aligned}$$

$$= N \sum_{l=1}^{q+1} \left( E \left[ \frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \cdot \sum_{k=l}^{q+1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho}) \right] - E \left[ \sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho}) \right] \right).$$

By a proof parallel with that of Theorem S2, we have:

$$E[\widehat{\delta}_{q+1}^{D'}(\boldsymbol{\Phi}, \boldsymbol{\rho})] = N \sum_{l=1}^{q+1} \left( \frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \cdot \sum_{k=l}^{q+1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) - \sum_{k=0}^{l-1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right)$$

Recall equation [S4] →

$$= \delta_{q+1}^D(\boldsymbol{\Phi}, \boldsymbol{\rho})$$

Note that under the null,  $\frac{\sum_{k=0}^{l-1} \rho_k}{\sum_{k=l}^{q+1} \rho_k} \hat{N}_{l-1}^u(\boldsymbol{\rho}) = \hat{N}_{l-1}^v(\boldsymbol{\rho})$  and equation (S15) can be written as

$$\sum_{l=1}^{q+1} \left[ \hat{N}_{l-1}^v(\boldsymbol{\rho}) \left( \hat{h}_l^u(\boldsymbol{\rho}) - \hat{h}_l^v(\boldsymbol{\rho}) \right) \right], \text{ which reduces to the hazard difference estimator in eAppendix}$$

6. Therefore, the same type of aggregated data required by the hazard difference estimator can be used by an unbiased estimator to quantify directly averted outcomes. However, when using aggregated data to identify the causal estimand in observational studies, we recognize that further procedures are needed to estimate  $\rho_x$  for  $x \in \{0, \dots, q+1\}$  because there are no random assignments of  $\mathbf{X}$ . We will further explore this in **eAppendix 13**.

## eAppendix 8. Parameters for simulations in the main text

**TABLE S4** | List of parameters for simulations

Parameter	Value	Definition
$\beta$	0.25	The number of effective contacts made by a typical infectious individual per day
$\mu$	varied	Infection-fatality rate
$\theta$	varied	1 – vaccine efficacy against infection (i.e., $VE_{inf}/100\%$ )
$\theta_1(t)$	1 if $t < 60$ days $\theta$ if $t \geq 60$ days	Same as above
$\kappa$	varied	1 – vaccine efficacy against death given infection (i.e., $VE_{death}/100\%$ )
$\kappa_1(t)$	1 if $t < 60$ days $\kappa$ if $t \geq 60$ days	Same as above
$\gamma$	0.07	Recovery rate per day

**TABLE S5** | List of initial conditions in the group under strategy  $\rho' = \{2,60; 0.2,0.3\}$

Variable	Initial condition(s)	Definition
$N(0)$	300,000	Number alive at baseline
$S_2(0)$	$(300,000 - 300) * (1 - 0.2 - 0.3)$	Number of susceptible individuals randomized to receive no vaccination
$S_1(0)$	$(300,000 - 300) * 0.3$	Number of susceptible individuals randomized to receive on Day 60
$S_0(0)$	$(300,000 - 300) * 0.2$	Number of susceptible individuals randomized to receive vaccination at baseline
$I_2(0)$	$300 * (1 - 0.2 - 0.3)$	Number of infectious individuals randomized to receive no vaccination
$I_1(0)$	$300 * 0.3$	Number of infectious individuals randomized to receive vaccination on Day 60
$I_0(0)$	$300 * 0.2$	Number of infectious individuals randomized to receive vaccination at baseline
$R_2(0)$	0	Number of recovered individuals randomized to receive no vaccination
$R_1(0)$	0	Number of recovered individuals randomized to receive vaccination on Day 60

$R_0(0)$	0	Number of recovered individuals randomized to receive vaccination at baseline
$D_2(0)$	0	Number of individuals randomized to receive no vaccination who died due to infection
$D_1(0)$	0	Number of individuals randomized to receive vaccination on Day 60 who died due to infection
$D_0(0)$	0	Number of individuals randomized to receive vaccination at baseline who died due to infection

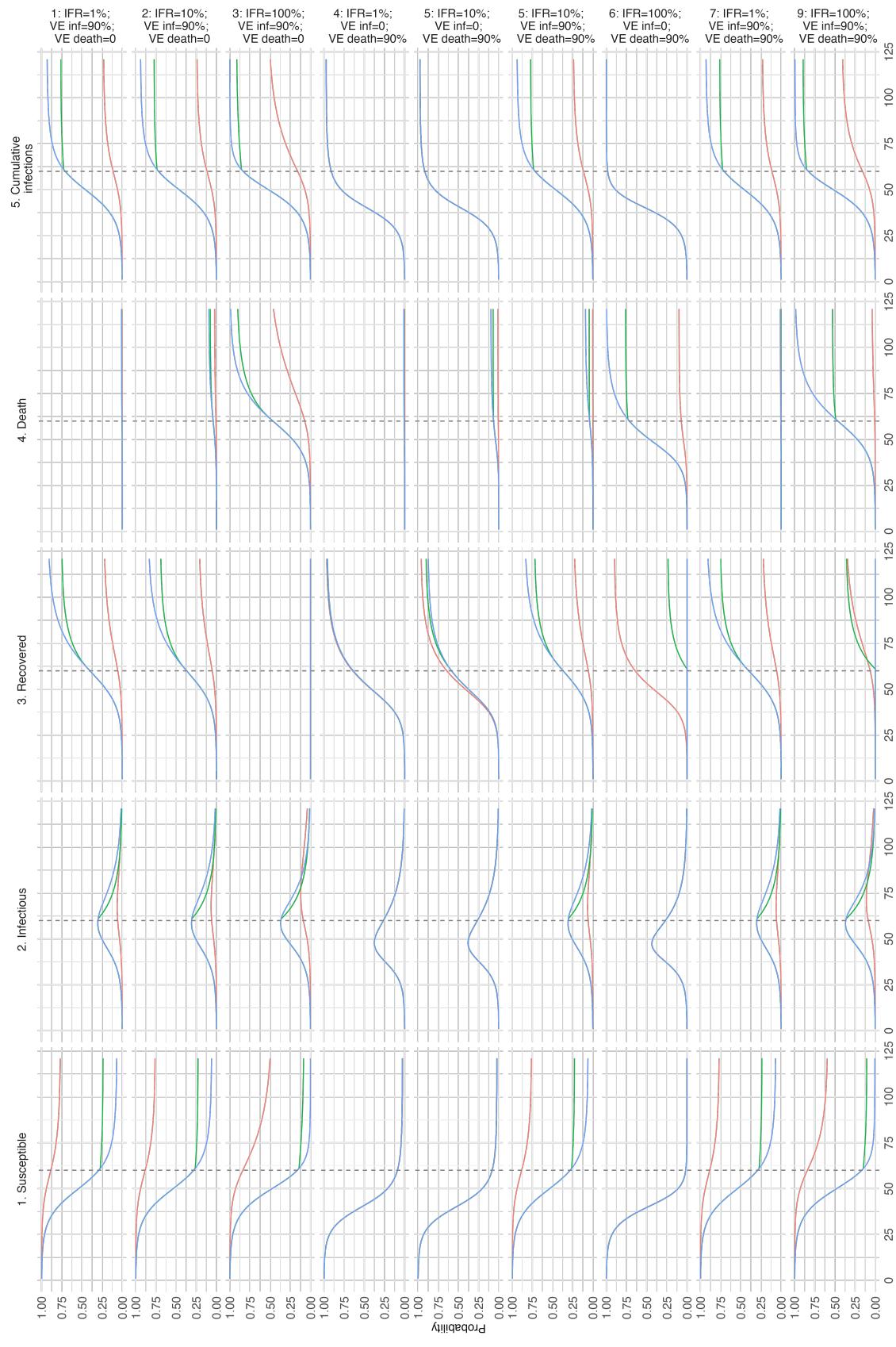
**eAppendix 9. Bias of the hazard difference estimator relative to the causal estimand on the absolute and relative scales**

**TABLE S6** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand (reference) for averted outcomes under strategy  $\rho'$  under scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ ), as outlined in Table 1.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9
IFR: 1%, $VE_{inf}$ = 90%, $VE_{death}$ = 0%	IFR: 10%, $VE_{inf}$ = 90%, $VE_{death}$ = 0%	IFR: 100%, $VE_{inf}$ = 0%, $VE_{death}$ = 90%	IFR: 10%, $VE_{inf}$ = 0%, $VE_{death}$ = 90%	IFR: 100%, $VE_{inf}$ = 0%, $VE_{death}$ = 90%	IFR: 100%, $VE_{inf}$ = 0%, $VE_{death}$ = 90%	IFR: 1%, $VE_{inf}$ = 90%, $VE_{death}$ = 90%	IFR: 10%, $VE_{inf}$ = 90%, $VE_{death}$ = 90%	IFR: 100%, $VE_{inf}$ = 90%, $VE_{death}$ = 90%	IFR: 100%, $VE_{inf}$ = 90%, $VE_{death}$ = 90%
<b>Averted infections <sup>a</sup></b>	27139.84 (47.39%)	28279.75 (49.63%)	40903.49 (108.29%)	0.00	0.00	0.00	27127.71 (47.37%)	28151.4 (49.43%)	39707.74 (89.09%)
<b>Averted deaths <sup>a</sup></b>	1.08 (0.19%)	114.18 (2.04%)	24399.47 (61.16%)	1.09 (0.14%)	116.4 (1.52%)	38964.52 (51.95%)	1.24 (0.13%)	131.14 (1.38%)	28154.09 (28.75%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places.

## eAppendix 10. Disease dynamics under each scenario



**FIGURE S3 |** Disease dynamics under strategy  $\boldsymbol{\rho}' = \{2, 60; 0.2, 0.3\}$  in scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ ). Red line represents individuals assigned to vaccination at baseline, green for those assigned to vaccination at Day 60, and blue for those remaining unvaccinated throughout. Grey dashed line indicates Day 60.

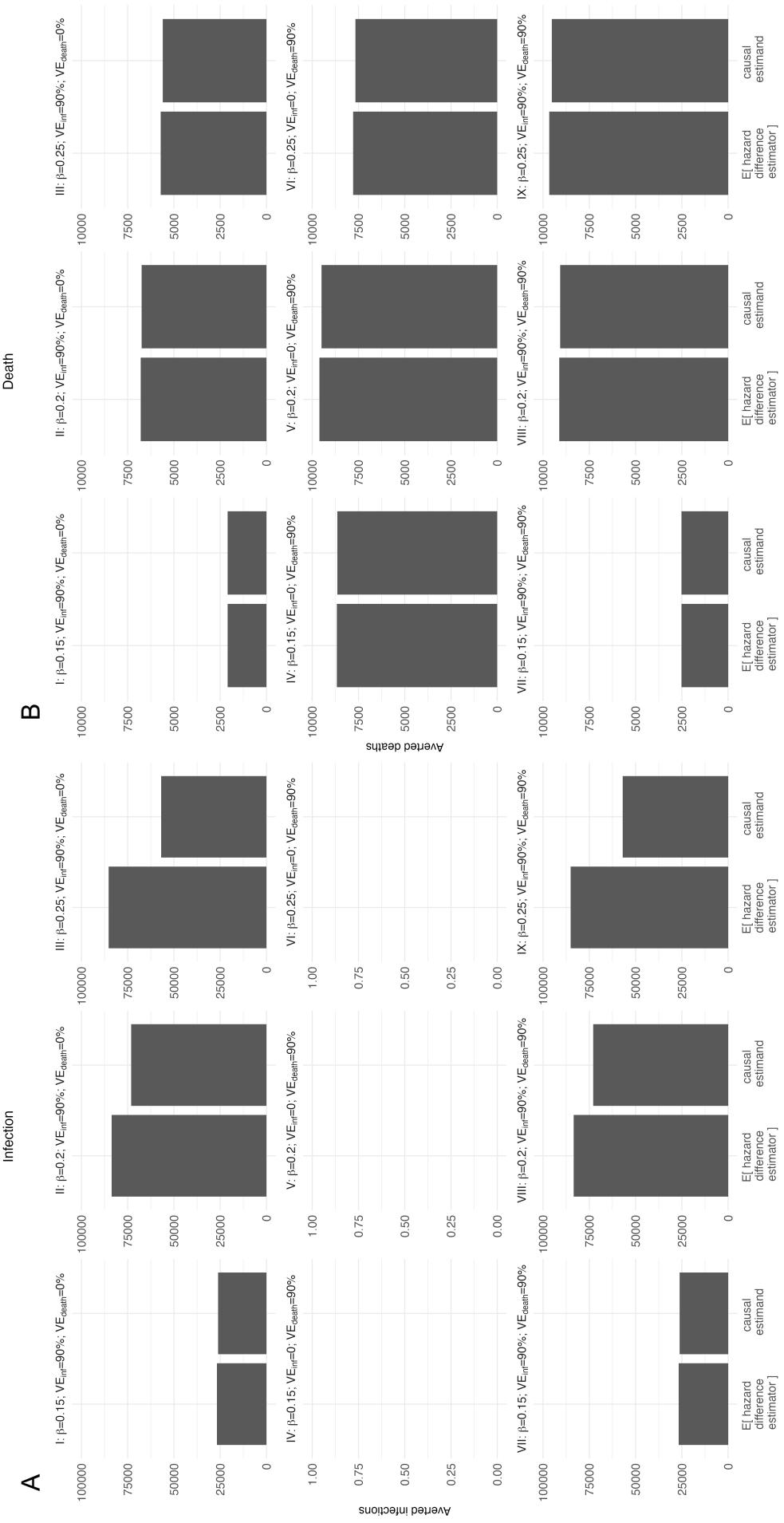
## eAppendix 11. Scenarios for simulations by varying the number of effective contacts

In **Figure 2B**, we varied IFR and showed that  $\widehat{\delta}_2^{D^*}(\boldsymbol{\phi}, \boldsymbol{\rho}')$  substantially overestimates averted deaths when  $IFR = 100\%$ , due to the stark difference in relative survival between vaccinated and not-yet-vaccinated individuals. However, when  $IFR \leq 10\%$ , the overestimation is trivial due to the similar survival across groups. Here, instead of varying IFR, we vary the number of effective contacts ( $\beta$ ) to explore the bias. Scenarios are specified in **Table S7**. All other parameters match the main analysis (**Table S4**), except for infection-fatality rate  $\mu$ , fixed at 0.1. Initial values also match the main analysis (**Table S5**).

**TABLE S7** | Scenarios varied by number of effective contacts ( $\beta$ ), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ )

Scenario	$\beta$	Vaccine efficacy
Scenario I	0.15	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario II	0.2	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario III	0.25	$VE_{inf} = 90\%; VE_{death} = 0\%$
Scenario IV	0.15	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario V	0.2	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario VI	0.25	$VE_{inf} = 0\%; VE_{death} = 90\%$
Scenario VII	0.15	$VE_{inf} = 90\%; VE_{death} = 90\%$
Scenario VIII	0.2	$VE_{inf} = 90\%; VE_{death} = 90\%$
Scenario IX	0.25	$VE_{inf} = 90\%; VE_{death} = 90\%$

**Figure S4** compares the expected value of hazard difference estimator with the causal estimand for averted infections and deaths in Scenarios as specified in **Table S7**, while **Table S8** shows the absolute and percentage bias. As shown in **Figure S4** and **Table S8**, the hazard difference estimator overestimates averted infections in all Scenarios (except for averted infections when  $VE_{inf} = 0\%$ ; Scenarios IV to VI), and the bias increases with  $\beta$ . This is because susceptibles (from infections) are preferentially depleted among unvaccinated individuals at a faster rate when  $\beta$  is higher and  $VE_{inf} = 90\%$ .



**FIGURE S4 |** Infections (A) and deaths (B) directly averted by strategy  $\rho' = \{2,60;0,2,0,3\}$  among vaccinated individuals under different scenarios varied by number of effective contacts ( $\beta$ ), vaccine efficacy against infection ( $VE_{inf}$ ), and vaccine efficacy against death given infection ( $VE_{death}$ ), as specified in **Table S7**

**TABLE S8** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand (reference) for averted outcomes under strategy  $\mathbf{p}' = \{2, 60; 0.2, 0.3\}$  under scenarios varied by the number of effective contacts ( $\beta$ ), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ ) as specified in **Table S7**

	Scenario I	Scenario II	Scenario III	Scenario IV	Scenario V	Scenario VI	Scenario VII	Scenario VIII	Scenario IX
	$\beta = 0.15$ , $VE_{inf} = 90\%$ , $VE_{death} = 0\%$	$\beta = 0.2$ , $VE_{inf} = 90\%$ , $VE_{death} = 0\%$	$\beta = 0.25$ , $VE_{inf} = 90\%$ , $VE_{death} = 0\%$	$\beta = 0.15$ , $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	$\beta = 0.2$ , $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	$\beta = 0.25$ , $VE_{inf} = 0\%$ , $VE_{death} = 90\%$	$\beta = 0.15$ , $VE_{inf} = 90\%$ , $VE_{death} = 90\%$	$\beta = 0.2$ , $VE_{inf} = 90\%$ , $VE_{death} = 90\%$	$\beta = 0.25$ , $VE_{inf} = 90\%$ , $VE_{death} = 90\%$
<b>Averted infections <sup>a</sup></b>	551.26 (2.1%)	10575.04 (14.47%)	28279.75 (49.63%)	0.00	0.00	0.00	550.01 (2.1%)	10526.5 (14.44%)	28151.4 (49.43%)
<b>Averted deaths <sup>a</sup></b>	2.68 (0.13%)	45.64 (0.68%)	114.18 (2.04%)	25.47 (0.29%)	112.22 (1.18%)	116.4 (1.52%)	3.03 (0.12%)	50.86 (0.56%)	131.14 (1.38%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places.

## eAppendix 12. Simulations with more realistic parameter values

In the main text, we considered extreme combinations of IFR,  $VE_{inf}$ , and  $VE_{death}$  for illustrative purposes. Here, we consider more realistic parameter values estimated by previous studies on seasonal flu, measles, and COVID-19 (wild-type strain), as shown in **Table S9**. As in the main text, we simulate strategy  $\rho' = \{2,60; 0.2,0.3\}$ , using the same model structure as in equation (8). Our objective is not to realistically model the dynamics of these epidemics—which would involve population heterogeneity, more complex infection characteristics (i.e., breakthrough infections, re-infections), waning immunity, vaccination at more time steps, and parameters that vary by time and context. Instead, we simply aim to test the robustness of our results across a broader and more realistic range of parameter values compared to those parameters used in the main text (as presented in **Table S4**).

**TABLE S9** | More realistic parameter values for seasonal flu, measles, and COVID-19 (wildtype). Definitions for the parameters are provided in **Table S4**. Model equation is shown in equation (8).

Parameter	Seasonal flu	Measles	COVID-19
$\beta^a$	0.41 <sup>15</sup>	1.26 <sup>16</sup>	0.15 <sup>17</sup>
$\mu$	3% <sup>18</sup>	1.3% <sup>19</sup>	0.11% <sup>20</sup>
$\theta$	0.66 <sup>21</sup>	0.05 <sup>22</sup>	0.05 <sup>23</sup>
$\kappa$	0.69 <sup>24</sup>	0 <sup>25</sup>	0.04 <sup>26</sup>
$\gamma$	0.21 <sup>15</sup>	0.07 <sup>27</sup>	0.07 <sup>28</sup>

<sup>a</sup>  $\beta$  is obtained from applying the formula  $\beta = R_0 \cdot \gamma$ , where  $R_0$  = basic reproduction number.

As **Figure S5** and **Table S10** shows, the hazard difference estimator substantially overestimates averted infections for measles due to high basic reproduction number ( $R_0=18$ ) and high  $VE_{inf}$ . It also slightly overestimates averted infections for seasonal flu and COVID-19 (wild-type), given a low  $R_0$  (i.e.,  $R_0 \leq 2.2$ ) and  $VE_{inf} > 0$ . The overestimation of averted deaths is trivial due to the low IFR ( $\leq 3\%$  for all pathogens).

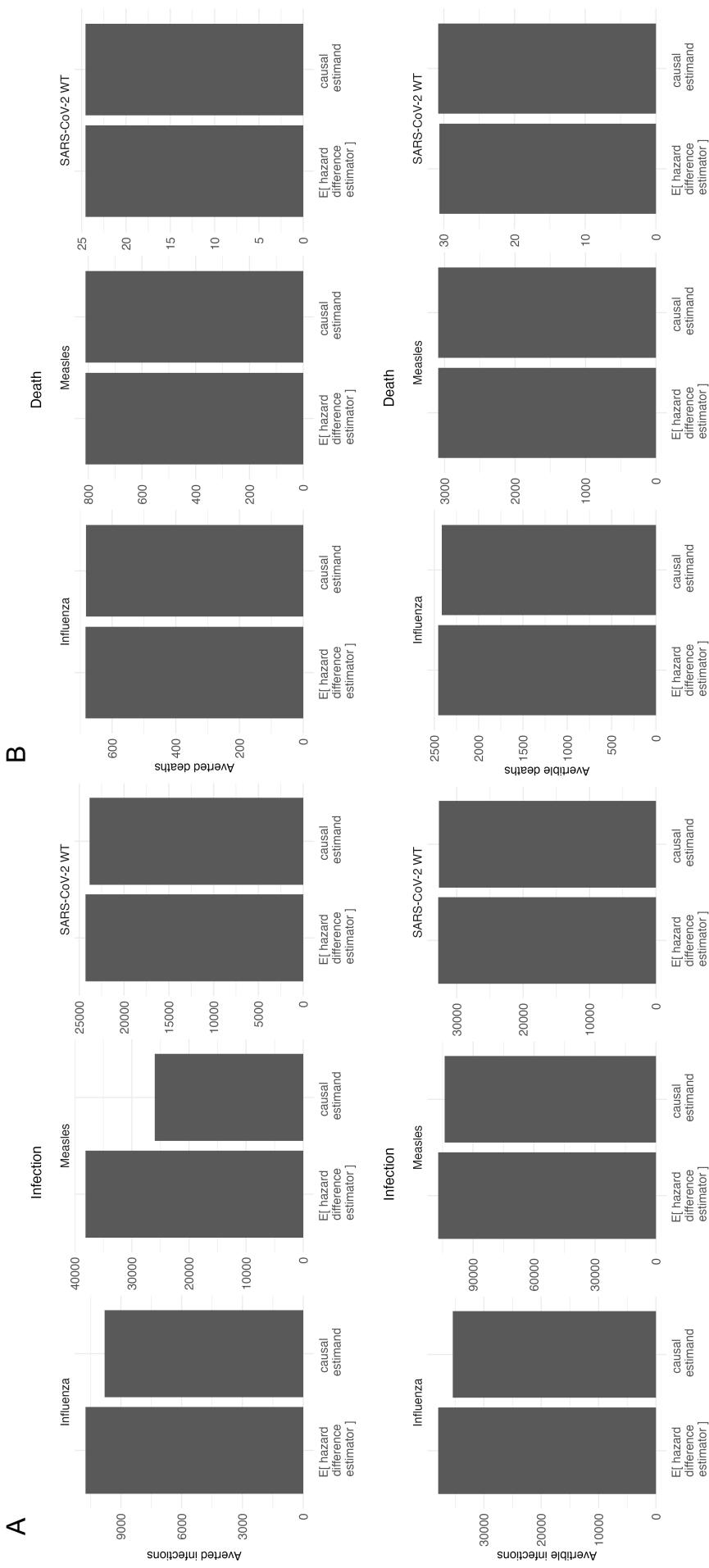
For avertible outcomes (defined in **eAppendix 14**), the hazard difference estimator slightly underestimates avertible deaths for COVID-19. As shown later, this occurs because the hazard

difference estimator neglects the term  $\rho_1 \cdot [\Delta\bar{Y}_2(1; \mathbf{p}) - \Delta\bar{Y}_2(0; \mathbf{p})]$  in the causal estimand (equation [S20]). This term is positive for COVID-19 (i.e., the period incidence of death for individuals with  $x = 1$  is higher than that for individuals with  $x = 0$ ), such that hazard difference estimator underestimates the causal estimand. For measles, there is no bias for avertible deaths under the hazard difference estimator because  $\Delta\bar{Y}_2(0; \mathbf{p}) = \Delta\bar{Y}_2(1; \mathbf{p}) = 0$  (See equation [S20] and compare it with equation [S24]).

**TABLE S10** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand (reference) for averted and avertable outcomes under strategy  $\mathbf{p}' = (2, 60; 0.2, 0.3)$  under more realistic parameter values for seasonal flu, measles, and COVID-19 (wild type strain)

	<b>Seasonal flu</b>	<b>Measles</b>	<b>COVID-19 (wild type strain)</b>
<b>Averted infections<sup>a</sup></b>	939.98 (9.59%)	12133.48 (46.69%)	462.12 (1.94%)
<b>Averted deaths<sup>a</sup></b>	1.02 (0.15%)	0.26 (0.03%)	0.00 (0.00%)
<b>Avertible infections<sup>a</sup></b>	2524.04 (7.13%)	3153.28 (3.03%)	99.9 (0.31%)
<b>Avertible deaths<sup>a</sup></b>	39.49 (1.63%)	0 (0%)	-0.16 (-0.53%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places.



**FIGURE S5 |** Infections (A) and deaths (B) directly averted by strategy  $\rho' = \{2,60; 0,2,0,3\}$  among vaccinated individuals under more realistic parameter values for seasonal flu, measles, and COVID-19 (wild type strain)

### eAppendix 13. Using observational data aggregated by vaccination status for estimation

eAppendix 7 defines an alternative unbiased estimator (equation [S15]) using RCT data aggregated by vaccination status. However, such an RCT is often not feasible in reality, and here we show how to identify the causal estimand using observational data aggregated by vaccination status (i.e., similar data structure as the national vaccine system<sup>29</sup>).

#### 1 Obtaining $\hat{\rho}_x$ using observational data

In the proposed RCT, we specify the strategy  $\boldsymbol{\rho}$  and the proportions  $\{\rho_0, \dots, \rho_{q+1}\}$  by design. In an observational study, the investigators make no assignments, rather we only observe the time at which individuals are vaccinated (denoted as  $\mathbf{X}^*$ ). In this case, the proportions  $\rho_x$  for  $x \in \{0, \dots, q+1\}$ , must be estimated from the data. However, we cannot assume  $\rho_x = E \left[ \frac{\sum_{j=1}^N I[X_j^* = x^*]}{N} \right]$  for  $x^* \in \{0, \dots, q+1\}$  due to immortal time bias (i.e., since individuals who die before interval 1 cannot receive vaccination, individuals with  $x^* = 1$  are by definition “immortal” before interval 1).<sup>30</sup> Here, we discuss the procedures to estimate  $\rho_x$  for  $x \in \{0, \dots, q+1\}$ , addressing the immortal time bias. Note we assume the absence of confounding and additional selection bias at baseline and throughout follow-up. These assumptions are implausible in observational studies, but addressing all their violations is beyond the scope of this paper. Instead, this section focuses on identifying the causal estimand using observational data aggregated by vaccination status, addressing the immortal time bias that would otherwise exist even in the absence of confounding or additional selection bias.

Without loss of generality, consider discrete-time outcome  $Y$  for *death*. The SIRD-model in equation (S14) can be simplified into the following equation for discrete-time outcome  $Y$ :

$$N_{l+1}^v(\boldsymbol{\rho}) = N_l^v(\boldsymbol{\rho}) \cdot (1 - h_{l+1}^v(\boldsymbol{\rho})) + N \cdot \rho_{l+1} \cdot \prod_{k=0}^{l+1} (1 - h_k^u(\boldsymbol{\rho})) \quad (\text{S16})$$

where  $N_l^v(\mathbf{p}) = N \sum_{k=0}^l \rho_k (1 - \bar{Y}_l(k; \mathbf{p}))$  is the number survived among vaccinated individuals by interval  $l$ , and  $h_{l+1}^v(\mathbf{p}) = \frac{\sum_{k=0}^l \rho_k \Delta \bar{Y}_{l+1}(k; \mathbf{p})}{\sum_{k=0}^l \rho_k (1 - \bar{Y}_l(k; \mathbf{p}))}$  (or  $h_{l+1}^u(\mathbf{p}) = \frac{\sum_{k=l+1}^{q+1} \rho_k \Delta \bar{Y}_{l+1}(k; \mathbf{p})}{\sum_{k=l+1}^{q+1} \rho_k (1 - \bar{Y}_l(k; \mathbf{p}))}$ ) is the hazard from interval  $l$  to  $l + 1$  among vaccinated (or unvaccinated) individuals.

Therefore,  $\rho_{l+1}$  can be estimated by:

$$\hat{\rho}_{l+1} = \frac{\hat{N}_{l+1}^v(\mathbf{p}) - \hat{N}_l^v(\mathbf{p}) \cdot (1 - \hat{h}_{l+1}^v(\mathbf{p}))}{N \cdot \prod_{k=0}^{l+1} (1 - \hat{h}_k^u(\mathbf{p}))}. \quad (\text{S17})$$

To illustrate this, **Table S11** shows a simulated dataset that is aggregated by vaccination status in a hypothetical observational study in the absence of confounding or additional selection bias. The epidemic is simulated under the (unobserved) strategy  $\mathbf{p}' = \{2, d; 0.2, 0.3\}$  in Scenario 1 as described in **Table 1**.

**TABLE S11** | Dataset aggregated by vaccination status in a hypothetical observational study without confounding under unobserved strategy  $\mathbf{p}' = \{2, 60; 0.2, 0.3\}$

Interval ( $k$ )	Observed number survived among vaccinated individuals $(\hat{N}_k^v(\mathbf{p}'))^a$	Hazards of death among vaccinated individuals $(\hat{h}_{k+1}^v(\mathbf{p}'))^b$	Observed number survived among not-yet-vaccinated individuals $(\hat{N}_k^u(\mathbf{p}'))^a$	Hazards of death among not-yet- vaccinated individuals $(\hat{h}_{k+1}^u(\mathbf{p}'))^b$
0	60000	0.00060716	240000	0.00426391
1	149579.818	0.0026133	149360.413	0.00491036

<sup>a</sup>Number of individuals survived before the start of interval  $k$  among those who are (un)vaccinated at the start of the that interval.

<sup>b</sup>Hazard for events occurring after the start of interval  $k$  and before the start of interval  $k + 1$  (i.e., hazard leads survival by one interval).

Assuming that we know the baseline vaccination proportion  $\rho'_0 = 0.2$ , we are interested in recovering  $\rho'_1$  and  $\rho'_2$ . By substituting the data from **Table S11** into equation (S17), we have:

$$\hat{\rho}'_1 = \frac{\hat{N}_1^v(\mathbf{p}') - \hat{N}_0^v(\mathbf{p}') \cdot (1 - \hat{h}_1^v(\mathbf{p}'))}{\hat{N} \cdot (1 - \hat{h}_0^u(\mathbf{p}')) \cdot (1 - \hat{h}_1^u(\mathbf{p}'))} = \frac{149579.818 - 60000 \cdot (1 - 0.00060716)}{300000 \cdot 1 \cdot (1 - 0.00426391)} = 0.3 = \rho'_1$$

Then, we can easily obtain  $\hat{\rho}'_2 = 1 - 0.2 - 0.3 = 0.5$ .

### 3 Identifying causal estimand using observational data

Lastly, we verify that, the causal estimand can be identified using observational data aggregated by vaccination status (subject to immortal time bias until corrected by equation [S17]) in the absence of confounding or additional selection bias. By having  $\mathbf{p}' = \hat{\mathbf{p}}'$  and substituting the data from **Table S11** into equation (S15), we have:

$$\begin{aligned}
\widehat{\delta_{q+1}}^{D'}(\Phi, \mathbf{p}') &= \sum_{l=1}^2 \left( \frac{\sum_{k=0}^{l-1} \hat{\rho}'_k}{\sum_{k=l}^{q+1} \hat{\rho}'_k} \hat{N}_{l-1}^u(\mathbf{p}') \hat{h}_l^u(\mathbf{p}') - \hat{N}_{l-1}^v(\mathbf{p}') \hat{h}_l^v(\mathbf{p}') \right) \\
&= \left[ \frac{\hat{\rho}'_0}{\hat{\rho}'_1 + \hat{\rho}'_2} \hat{N}_0^u(\mathbf{p}') \hat{h}_1^u(\mathbf{p}') - \hat{N}_0^v(\mathbf{p}') \hat{h}_1^v(\mathbf{p}') \right] + \left[ \frac{\hat{\rho}'_0 + \hat{\rho}'_1}{\hat{\rho}'_2} \hat{N}_1^u(\mathbf{p}') \hat{h}_2^u(\mathbf{p}') - \hat{N}_1^v(\mathbf{p}') \hat{h}_2^v(\mathbf{p}') \right] \\
&= \left( \frac{0.2}{0.8} * 240000 * 0.00426391 - 60000.0 * 0.00060716 \right) + \left( \frac{0.5}{0.5} * 149360.413 * 0.00491036 - 149579.818 * 0.0026133 \right) \\
&= 561.92146
\end{aligned}$$

Consider the simulated data from a hypothetical RCT where we have cumulative incidence data disaggregated by vaccination time (**Table S12**). By substituting the RCT data into equation (S10) (i.e., the unbiased estimator), we estimate the averted deaths to be  $300000 * [0.2 * (0.00915333 - 0.00220958) + 0.3 * (0.00915333 - 0.00753893)] = 561.921$ , which is the same as the result obtained from **Table S11**. Therefore, in the absence of confounding or additional selection bias, we can use observational data aggregated by vaccination status to estimate cumulative incidence difference estimand for averted outcomes.

**TABLE S12** | Dataset disaggregated by vaccination time in a hypothetical one-stage RCT under strategy  $\mathbf{p}' = \{2,60; 0.2,0.3\}$

Vaccination time ( $x$ )	$\mathbf{p}'_x$	Cumulative incidence for death $\hat{Y}_2(x; \mathbf{p}')$
0	0.2	0.00220958
1	0.3	0.00753893
2	0.5	0.00915333

## eAppendix 14. Outcomes directly avertible by vaccination

In the main text, we define the causal estimand and estimators for quantifying outcomes directly averted by vaccination. Here, we propose the causal estimand and estimators for outcomes directly *avertible* by vaccination. The proposed causal estimand considers outcomes that could have been averted under full vaccination at baseline (denoted as  $\Psi = (q + 1, d; 1, \mathbf{0})$  for  $q \in \mathbb{N}$ ), but were not averted given the particular vaccination strategy  $\mathbf{p}$ .

### 1.1 Causal estimand for outcomes directly avertible by vaccination at a single time point

Our previous framework<sup>14</sup> defined the causal estimand for quantifying avertible outcomes for vaccination at a single time point—that is, outcomes that could have been directly averted under full vaccination at baseline  $\Psi = \{1, d; 1\}$  but were not averted given the particular strategy  $\mathbf{p} = \{1, d; \rho_0\}$ , where  $\rho_1 = 1 - \rho_0$ . In notation, the estimand is:

$$\delta_1^D(\Psi, \mathbf{p}) = N \cdot \rho_1 \cdot (\bar{Y}_1(1; \mathbf{p}) - \bar{Y}_1(0; \mathbf{p})) \quad (\text{S18})$$

### 1.2 Causal estimand for outcomes directly avertible by vaccination at two time points

Now, extending the estimand in equation (S18) for vaccination at two time points, we define the estimand to quantify outcomes that could have been directly averted under full vaccination at baseline  $\Psi = \{2, d; 1, 0\}$  but were not averted given the particular strategy  $\mathbf{p} = \{2, d; \rho_0, \rho_1\}$ . In notation, the estimand is:

$$\delta_2^D(\Psi, \mathbf{p}) = N \cdot [\rho_1(\bar{Y}_2(1; \mathbf{p}) - \bar{Y}_2(0; \mathbf{p})) + \rho_2(\bar{Y}_2(2; \mathbf{p}) - \bar{Y}_2(0; \mathbf{p}))] \quad (\text{S19})$$

$$\begin{aligned} &= N \cdot [(\rho_1 \Delta \bar{Y}_1(1; \mathbf{p}) + \rho_2 \Delta \bar{Y}_1(2; \mathbf{p})) - (\rho_1 + \rho_2) \cdot \Delta \bar{Y}_1(0; \mathbf{p}) + \rho_1 \cdot (\Delta \bar{Y}_2(1; \mathbf{p}) - \Delta \bar{Y}_2(0; \mathbf{p})) \\ &\quad + \rho_2 \cdot (\Delta \bar{Y}_2(2; \mathbf{p}) - \Delta \bar{Y}_2(0; \mathbf{p}))]. \quad (\text{S20}) \end{aligned}$$

In words, equation (S19) takes the difference in cumulative incidence between vaccination at baseline (i.e.,  $\bar{Y}_2(0; \mathbf{p})$ ) and vaccination at later time  $x \in \{1, 2\}$  (i.e.,  $\bar{Y}_2(1; \mathbf{p})$  or  $\bar{Y}_2(2; \mathbf{p})$ ).

multiplied with number of individuals assigned to  $x=1$  or  $2$  (i.e.,  $N \cdot \rho_1$  or  $N \cdot \rho_2$ ), respectively, and sum over  $x \in \{1,2\}$ .

### 1.3 Outcomes directly avertible by vaccination at an arbitrary number of vaccination times

Extending the estimand in equation (S19) for vaccination at an arbitrary number of vaccination times, we define the estimand to quantify outcomes that could have been directly averted under full vaccination at baseline  $\Psi = \{q + 1, d; 1, \mathbf{0}\}$  but were not averted given the particular strategy  $\boldsymbol{\rho} = \{q + 1, d; \rho_0, \dots, \rho_q\}$  for  $q \in \mathbb{N}$ . In notation, the estimand is:

$$\delta_{q+1}^D(\Psi, \boldsymbol{\rho}) = N \cdot \sum_{k=1}^{q+1} \rho_k \cdot (\bar{Y}_{q+1}(k; \boldsymbol{\rho}) - \bar{Y}_{q+1}(0; \boldsymbol{\rho})). \quad (\text{S21})$$

## 2 Unbiased estimator

Similar to the unbiased estimator for directly averted outcomes in equation (S10), directly avertible outcomes (i.e.,  $\delta_{q+1}^D(\Psi, \boldsymbol{\rho})$ ) can be identified by:

$$\widehat{\delta}_{q+1}^D(\Psi, \boldsymbol{\rho}) = N \cdot \sum_{k=1}^{q+1} \rho_k \cdot [\hat{Y}_{q+1}(k; \boldsymbol{\rho}) - \hat{Y}_{q+1}(0; \boldsymbol{\rho})] \quad (\text{S22})$$

where  $\hat{Y}_{q+1}(x; \boldsymbol{\rho}) = \frac{\sum_{j=1}^N Y_{q+1,j}(\mathbf{X}) I[X_j=x]}{\sum_{j=1}^N I[X_j=x]}$  for  $x \in \{0, \dots, q + 1\}$ . We know that  $\widehat{\delta}_{q+1}^D(\Psi, \boldsymbol{\rho})$  is

unbiased based on a proof parallel with the proof of Theorem S2.

### 3 Hazard difference estimator for outcomes directly avertible by vaccination

Some empirical studies<sup>31,32</sup> have used the hazard difference estimator to estimate vaccine-avertible outcomes, which takes the form:

$$\begin{aligned} \widehat{\delta}_{q+1}^{D*}(\Psi, \boldsymbol{\rho}) &= \sum_{l=1}^{q+1} [\hat{N}_{l-1}^u(\boldsymbol{\rho}) \cdot (\hat{h}_l^u(\boldsymbol{\rho}) - \hat{h}_l^v(\boldsymbol{\rho}))] \\ &= \hat{N}_0^u(\boldsymbol{\rho}) \cdot (\hat{h}_1^u(\boldsymbol{\rho}) - \hat{h}_1^v(\boldsymbol{\rho})) + \sum_{l=2}^{q+1} [\hat{N}_{l-1}^u(\boldsymbol{\rho}) \cdot (\hat{h}_l^u(\boldsymbol{\rho}) - \hat{h}_l^v(\boldsymbol{\rho}))] \end{aligned} \quad (\text{S23})$$

where  $\hat{N}_{l-1}^u(\boldsymbol{\rho}) = N \sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))$ ,  $\hat{h}_l^u(\boldsymbol{\rho}) = \frac{\sum_{k=l}^{q+1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}$ , and  $\hat{h}_l^v(\boldsymbol{\rho}) = \frac{\sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}$ . In words, the hazard difference estimator quantifies directly avertible outcomes by multiplying the hazard difference  $\hat{h}_l^u(\boldsymbol{\rho}) - \hat{h}_l^v(\boldsymbol{\rho})$  by the number survived among not-yet-vaccinated individuals  $\hat{N}_{l-1}^u(\boldsymbol{\rho})$ .

## 4 Bias of the hazard difference estimator relative to the causal estimand for vaccination at two time points

### 4.1 Analytic comparison

Consider vaccination at two time points. The hazard difference estimator for quantifying outcomes that could have been directly averted under full vaccination at baseline  $\boldsymbol{\Psi} = \{2, d; 1, 0\}$  but were not averted given the particular strategy  $\boldsymbol{\rho} = \{2, d; \rho_0, \rho_1\}$  is:

$$\begin{aligned}
\widehat{\delta}_2^{D^*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) &= \hat{N}_0^u(\boldsymbol{\rho}) (\hat{h}_1^u(\boldsymbol{\rho}) - \hat{h}_1^v(\boldsymbol{\rho})) + \hat{N}_1^u(\boldsymbol{\rho}) (\hat{h}_2^u(\boldsymbol{\rho}) - \hat{h}_2^v(\boldsymbol{\rho})) \\
&= N(\rho_1 + \rho_2) \left( \frac{\rho_1 \Delta \hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \hat{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta \hat{Y}_1(0; \boldsymbol{\rho}) \right) + N\rho_2 [1 - \hat{Y}_1(2; \boldsymbol{\rho})] \\
&\quad \cdot \left( \frac{\Delta \hat{Y}_2(2; \boldsymbol{\rho})}{1 - \hat{Y}_1(2; \boldsymbol{\rho})} - \frac{\rho_0 \Delta \hat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \hat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))} \right) \\
&= N \cdot \left[ (\rho_1 + \rho_2) \left( \frac{\rho_1 \Delta \hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \hat{Y}_1(2; \boldsymbol{\rho})}{\rho_1 + \rho_2} - \Delta \hat{Y}_1(0; \boldsymbol{\rho}) \right) + \rho_2 \right. \\
&\quad \cdot \left. \left( \Delta \hat{Y}_2(2; \boldsymbol{\rho}) - [1 - \hat{Y}_1(2; \boldsymbol{\rho})] \cdot \frac{\rho_0 \Delta \hat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \hat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))} \right) \right] \\
&= N \cdot \left[ \left( \rho_1 \Delta \hat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \hat{Y}_1(2; \boldsymbol{\rho}) - (\rho_1 + \rho_2) \cdot \Delta \hat{Y}_1(0; \boldsymbol{\rho}) \right) + \rho_2 \right. \\
&\quad \cdot \left. \left( \Delta \hat{Y}_2(2; \boldsymbol{\rho}) - \frac{(\rho_0 + \rho_1) \cdot (1 - \hat{Y}_1(2; \boldsymbol{\rho}))}{\rho_0 (1 - \hat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \hat{Y}_1(1; \boldsymbol{\rho}))} \cdot \frac{\rho_0 \Delta \hat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \hat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 + \rho_1} \right) \right].
\end{aligned}$$

Consider  $E [\widehat{\delta}_2^{D^*}(\boldsymbol{\Psi}, \boldsymbol{\rho})]$ .

$$\begin{aligned}
E \left[ \widehat{\delta}_2^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) \right] &= E \left[ N \right. \\
&\quad \cdot \left. \left[ \left( \rho_1 \Delta \widehat{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \widehat{Y}_1(2; \boldsymbol{\rho}) - (\rho_1 + \rho_2) \cdot \Delta \widehat{Y}_1(0; \boldsymbol{\rho}) \right) + \rho_2 \right. \\
&\quad \cdot \left. \left. \left( \Delta \widehat{Y}_2(2; \boldsymbol{\rho}) - \frac{(\rho_0 + \rho_1) \cdot (1 - \widehat{Y}_1(2; \boldsymbol{\rho}))}{\rho_0 (1 - \widehat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \widehat{Y}_1(1; \boldsymbol{\rho}))} \cdot \frac{\rho_0 \Delta \widehat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \widehat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 + \rho_1} \right) \right] \right] \\
&= N \cdot \left\{ \left( \rho_1 \Delta \bar{Y}_1(1; \boldsymbol{\rho}) + \rho_2 \Delta \bar{Y}_1(2; \boldsymbol{\rho}) \right) - (\rho_1 + \rho_2) \cdot \Delta \bar{Y}_1(0; \boldsymbol{\rho}) + \rho_2 \cdot \Delta \bar{Y}_2(2; \boldsymbol{\rho}) \right. \\
&\quad \left. - E \left[ \rho_2 \cdot \left( \frac{(\rho_0 + \rho_1) \cdot (1 - \widehat{Y}_1(2; \boldsymbol{\rho}))}{\rho_0 (1 - \widehat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \widehat{Y}_1(1; \boldsymbol{\rho}))} \cdot \frac{\rho_0 \Delta \widehat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \widehat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 + \rho_1} \right) \right] \right\} \quad (\text{S24})
\end{aligned}$$

Comparing equation (S20) with equation (S24), we know that  $E \left[ \widehat{\delta}_2^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) \right] \neq \delta_2^D(\boldsymbol{\Psi}, \boldsymbol{\rho})$  if vaccine has a non-null effect, as equation (S24) misses the term  $\rho_1 \cdot [\Delta \bar{Y}_2(1; \boldsymbol{\rho}) - \Delta \bar{Y}_2(0; \boldsymbol{\rho})]$  and

$$E \left[ \rho_2 \cdot \frac{(\rho_0 + \rho_1) \cdot (1 - \widehat{Y}_1(2; \boldsymbol{\rho}))}{\rho_0 (1 - \widehat{Y}_1(0; \boldsymbol{\rho})) + \rho_1 (1 - \widehat{Y}_1(1; \boldsymbol{\rho}))} \cdot \frac{\rho_0 \Delta \widehat{Y}_2(0; \boldsymbol{\rho}) + \rho_1 \Delta \widehat{Y}_2(1; \boldsymbol{\rho})}{\rho_0 + \rho_1} \right] \neq \rho_2 \cdot \Delta \bar{Y}_2(0; \boldsymbol{\rho}).$$

## 4.2 Simulation

We simulate the epidemic under strategy  $\boldsymbol{\rho}' = (2, 60; 0.2, 0.3)$  based on the same model (equation [8]), scenarios (**Table 1**), parameters (**Table S4**), and initial conditions (**Table S5**) as with the simulations in the main text. We examine the bias of the hazard difference estimator  $\widehat{\delta}_2^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}')$  relative to the causal estimand  $\delta_2^D(\boldsymbol{\Psi}, \boldsymbol{\rho}')$ , where  $\boldsymbol{\Psi} = \{2, 60; 1, 0\}$ , and identify the conditions under which the bias would be substantial.

The hazard difference estimator overestimates directly avertible infections when  $VE_{inf} = 90\%$  (**Figure S6A; Table S13**). However, the hazard difference estimator overestimates directly avertible deaths in Scenarios 3 to 6 and underestimates directly avertible deaths in other scenarios (**Figure S6B; Table S13**).

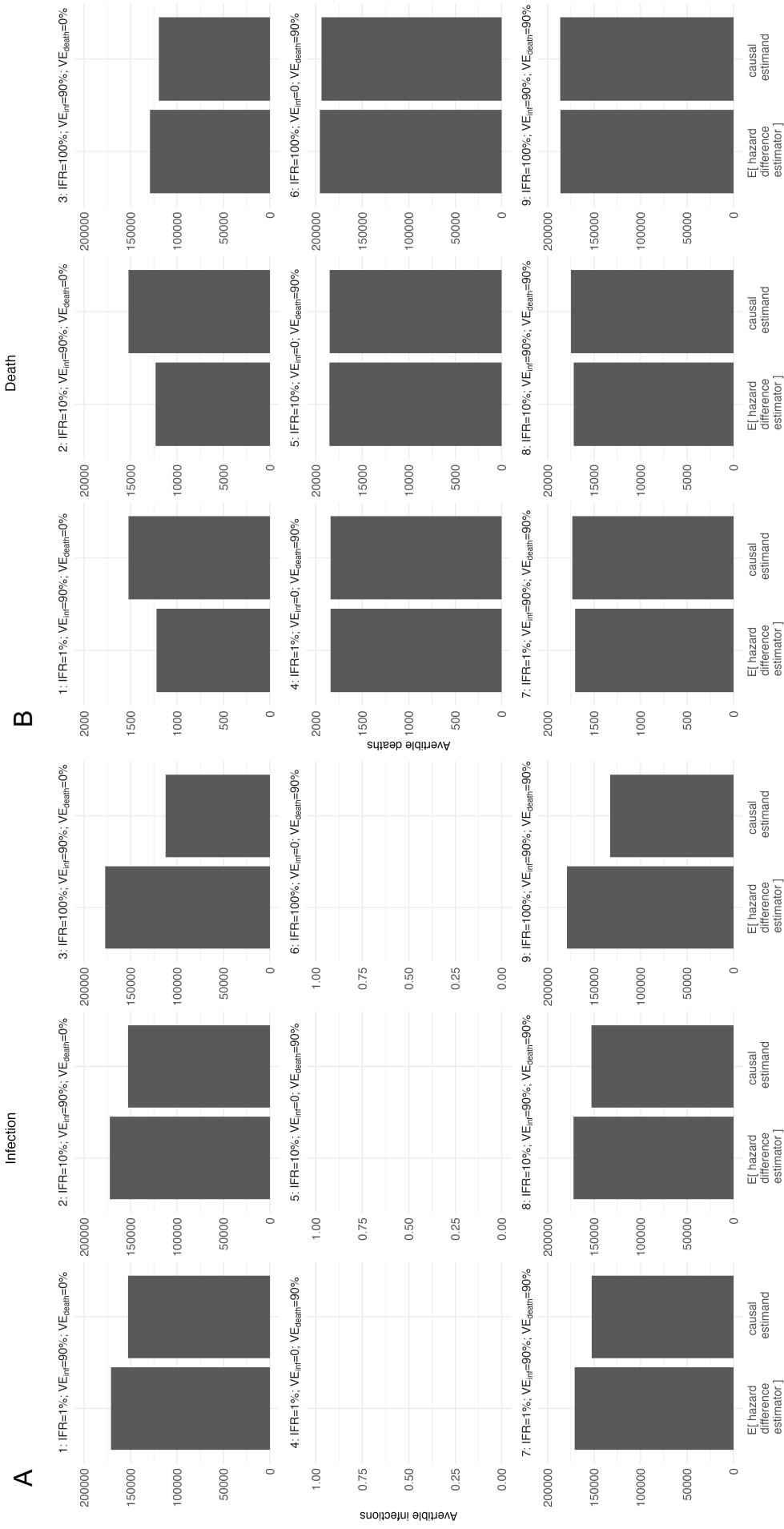
Take Scenario 3 where IFR=100% as an example. The hazard difference estimator overestimates avertible deaths. Due to high  $VE_{inf}$ , survival is lower among  $x = 2$  compared to the average of  $x = 1$  or  $x = 0$  (**Figure S3**). Specifically,  $\frac{(\rho'_0 + \rho'_1) \cdot (1 - \hat{Y}_1(2; \mathbf{p}'))}{\rho'_0(1 - \hat{Y}_1(0; \mathbf{p}')) + \rho'_1(1 - \hat{Y}_1(1; \mathbf{p}'))} < 1$  in equation (S24), such that  $E \left[ \rho'_2 \cdot \frac{(\rho'_0 + \rho'_1) \cdot (1 - \hat{Y}_1(2; \mathbf{p}'))}{\rho'_0(1 - \hat{Y}_1(0; \mathbf{p}')) + \rho'_1(1 - \hat{Y}_1(1; \mathbf{p}'))} \cdot \frac{\rho'_0 \Delta \hat{Y}_2(0; \mathbf{p}') + \rho'_1 \Delta \hat{Y}_2(1; \mathbf{p}')}{\rho'_0 + \rho'_1} \right]$  in equation (S24) is smaller than  $\rho'_2 \cdot \Delta \bar{Y}_2(0; \mathbf{p}')$  in equation (S20).

However, in Scenario 2 where IFR=10%, the hazard difference estimator underestimates avertible deaths because the relative survival is similar (i.e.,  $\frac{(\rho'_0 + \rho'_1) \cdot (1 - \hat{Y}_1(2; \mathbf{p}'))}{\rho'_0(1 - \hat{Y}_1(0; \mathbf{p}')) + \rho'_1(1 - \hat{Y}_1(1; \mathbf{p}'))} \approx 1$ ), such that  $E \left[ \rho'_2 \cdot \frac{(\rho'_0 + \rho'_1) \cdot (1 - \hat{Y}_1(2; \mathbf{p}'))}{\rho'_0(1 - \hat{Y}_1(0; \mathbf{p}')) + \rho'_1(1 - \hat{Y}_1(1; \mathbf{p}'))} \cdot \frac{\rho'_0 \Delta \hat{Y}_2(0; \mathbf{p}') + \rho'_1 \Delta \hat{Y}_2(1; \mathbf{p}')}{\rho'_0 + \rho'_1} \right]$  in equation (S24) is greater than  $\rho'_2 \cdot \Delta \bar{Y}_2(0; \mathbf{p}')$  in equation (S20).

**TABLE S13** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand (reference) for avertible outcomes under strategy  $\boldsymbol{\rho}' = (2.60; 0.2, 0.3)$  in scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ )

	<b>Scenario 1</b>	<b>Scenario 2</b>	<b>Scenario 3</b>	<b>Scenario 4</b>	<b>Scenario 5</b>	<b>Scenario 6</b>	<b>Scenario 7</b>	<b>Scenario 8</b>	<b>Scenario 9</b>
IFR: 1%, $VE_{inf} =$ 90%, $VE_{death} =$ 0%	IFR: 10%, $VE_{inf} =$ 90%, $VE_{death} =$ 0%	IFR: 100%, $VE_{inf} =$ 90%, $VE_{death} =$ 0%	IFR: 1%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 10%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 100%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 1%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 10%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 100%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%	IFR: 100%, $VE_{inf} = 0\%$ , $VE_{death} =$ 90%
<b>Avertible infections <sup>a</sup></b>	18348.79 (12.02%)	19706.57 (12.92%)	65014.35 (57.86%)	0.00	0.00	0.00	18333.15 (12.01%)	19528.3 (12.79%)	46502.42 (35.06%)
<b>Avertible deaths <sup>a</sup></b>	-300.49 (-19.75%)	-2911.27 (-19.12%)	9526.54 (7.96%)	0.11 (0.01%)	11.35 (0.06%)	1957.46 (1.01%)	-30.05 (-1.73%)	-291.24 (-1.67%)	-132.51 (-0.07%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places.



**FIGURE S6 |** Infections (A) and deaths (B) directly avertible among unvaccinated individuals with strategy  $\mathbf{p}' = (2, 60; 0.2, 0.3)$  under different scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ )

## 5 Bias of the hazard difference estimator relative to the causal estimand for vaccination at an arbitrary number of vaccination times

### 5.1 Analytic comparison

Now consider vaccination at an arbitrary number of vaccination times (i.e.,  $q \in \mathbb{N}$ ). First, expand

$\delta_{q+1}^D(\Psi, \rho)$  in equation (S21) as:

$$\begin{aligned}
\delta_{q+1}^D(\Psi, \rho) &= N \cdot \sum_{k=1}^{q+1} \rho_k \cdot (\bar{Y}_{q+1}(k; \rho) - \bar{Y}_{q+1}(0; \rho)) \\
&= N \cdot \sum_{k=1}^{q+1} \rho_k \cdot \left( \sum_{l=1}^{q+1} \Delta \bar{Y}_l(k; \rho) - \sum_{l=1}^{q+1} \Delta \bar{Y}_l(0; \rho) \right) \\
&= N \cdot \sum_{k=1}^{q+1} \rho_k \cdot \left( \Delta \bar{Y}_1(k; \rho) + \sum_{l=2}^{q+1} \Delta \bar{Y}_l(k; \rho) - \Delta \bar{Y}_1(0; \rho) - \sum_{l=2}^{q+1} \Delta \bar{Y}_l(0; \rho) \right) \\
&= N \cdot \left[ \sum_{k=1}^{q+1} \rho_k \cdot (\Delta \bar{Y}_1(k; \rho) - \Delta \bar{Y}_1(0; \rho)) + \sum_{k=1}^{q+1} \rho_k \cdot \left( \sum_{l=2}^{q+1} \Delta \bar{Y}_l(k; \rho) - \sum_{l=2}^{q+1} \Delta \bar{Y}_l(0; \rho) \right) \right] \quad (\text{S25})
\end{aligned}$$

Then, expand  $\widehat{\delta}_{q+1}^{D^*}(\Psi, \rho)$  in equation (S23) as:

$$\begin{aligned}
\widehat{\delta}_{q+1}^{D^*}(\Psi, \rho) &= \sum_{l=1}^{q+1} \left[ \widehat{N}_{l-1}^u(\rho) \cdot (\widehat{h}_l^u(\rho) - \widehat{h}_l^v(\rho)) \right] \\
&= \sum_{l=1}^{q+1} \left\{ N \left[ \sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho)) \right] \cdot \left[ \frac{\sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \rho)}{\sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho))} - \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \rho)}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho))} \right] \right\} \\
&= N \sum_{l=1}^{q+1} \left\{ \sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \rho) - \left[ \sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho)) \right] \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \rho)}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho))} \right\}
\end{aligned}$$

Pulling  $l = 1$  from the summation  $\rightarrow$

$$\begin{aligned}
&= N \left\{ \sum_{k=1}^{q+1} \rho_k \Delta \widehat{Y}_1(k; \rho) - \left( \sum_{k=1}^{q+1} \rho_k \right) \cdot \Delta \widehat{Y}_1(0; \rho) \right. \\
&\quad \left. + \sum_{l=2}^{q+1} \left\{ \sum_{k=l}^{q+1} \rho_k \Delta \widehat{Y}_l(k; \rho) - \left[ \sum_{k=l}^{q+1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho)) \right] \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \widehat{Y}_l(k; \rho)}{\sum_{k=0}^{l-1} \rho_k (1 - \widehat{Y}_{l-1}(k; \rho))} \right\} \right\}
\end{aligned}$$

$$= N \left\{ \sum_{k=1}^{q+1} \rho_k (\Delta \hat{Y}_1(k; \boldsymbol{\rho}) - \Delta \hat{Y}_1(0; \boldsymbol{\rho})) + \sum_{l=2}^{q+1} \left[ \sum_{k=l}^{q+1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho}) - \left( \sum_{k=l}^{q+1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))} \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k} \right] \right\}.$$

Now consider  $E \left[ \widehat{\delta_{q+1}}^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) \right]$ .

$$E \left[ \widehat{\delta_{q+1}}^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) \right] = E \left[ N \left\{ \sum_{k=1}^{q+1} \rho_k (\Delta \hat{Y}_1(k; \boldsymbol{\rho}) - \Delta \hat{Y}_1(0; \boldsymbol{\rho})) + \sum_{l=2}^{q+1} \left[ \sum_{k=l}^{q+1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho}) - \left( \sum_{k=l}^{q+1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))} \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k} \right] \right\} \right]$$

Using arguments parallel with those in the proof of Theorem S2 →

$$= N \cdot \left\{ \left[ \sum_{k=1}^{q+1} \rho_k \cdot (\Delta \bar{Y}_1(k; \boldsymbol{\rho}) - \Delta \bar{Y}_1(0; \boldsymbol{\rho})) \right] + \sum_{l=2}^{q+1} \left[ \sum_{k=l}^{q+1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho}) \right] - E \left[ \sum_{l=2}^{q+1} \left[ \left( \sum_{k=l}^{q+1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))} \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k} \right] \right] \right\} \quad (\text{S26})$$

Equation (S25) and equation (S26) have the same first term. However, if vaccination has a non-null effect,  $\sum_{k=1}^{q+1} \rho_k \cdot (\sum_{l=2}^{q+1} \Delta \bar{Y}_l(k; \boldsymbol{\rho}) - \sum_{l=2}^{q+1} \Delta \bar{Y}_l(0; \boldsymbol{\rho})) = \sum_{k=1}^{q+1} \rho_k \cdot \sum_{l=2}^{q+1} \Delta \bar{Y}_l(k; \boldsymbol{\rho}) - \sum_{k=1}^{q+1} \rho_k \cdot \sum_{l=2}^{q+1} \Delta \bar{Y}_l(0; \boldsymbol{\rho})$  in equation (S25) does not equal to  $\sum_{l=2}^{q+1} [\sum_{k=l}^{q+1} \rho_k \Delta \bar{Y}_l(k; \boldsymbol{\rho})] - E \left[ \sum_{l=2}^{q+1} \left[ \left( \sum_{k=l}^{q+1} \rho_k \right) \cdot \frac{\sum_{k=l}^{q+1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))}{\sum_{k=0}^{l-1} \rho_k (1 - \hat{Y}_{l-1}(k; \boldsymbol{\rho}))} \cdot \frac{\sum_{k=0}^{l-1} \rho_k \Delta \hat{Y}_l(k; \boldsymbol{\rho})}{\sum_{k=0}^{l-1} \rho_k} \right] \right]$  in equation (S26), implying  $E \left[ \widehat{\delta_{q+1}}^{D*}(\boldsymbol{\Psi}, \boldsymbol{\rho}) \right] \neq \delta_{q+1}^D(\boldsymbol{\Psi}, \boldsymbol{\rho})$ .

## 5.2 Simulation

We simulate the epidemic under strategy  $\mathbf{p}'' = \{99.7; \mathbf{0.01}\}$  based on the model in equation (S14), scenarios in **Table 1**, initial values in **Table S2** and parameters in **Table S4**. Similar to findings observed under strategy  $\mathbf{p}'$ , the hazard difference estimator overestimates directly avertible infections when  $VE_{inf}=90\%$  (**Figure S7A, Table S14**). It substantially overestimates directly avertible deaths when  $IFR=100\%$  and  $VE_{death} = 0\%$  (Scenario 3), slightly overestimates them in Scenarios 4, 5, 6 and 9, and underestimates them in Scenarios 1, 2, 7, and 8 (**Figure S7B, Table S14**).

**TABLE S14** | Absolute and percentage bias of the hazard difference estimator relative to the causal estimand for avertible outcomes under strategy  $\rho'' = \{99.7; \mathbf{0.1}\}$  in scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection (VE<sub>inf</sub>) and vaccine efficacy against death given infection (VE<sub>death</sub>)

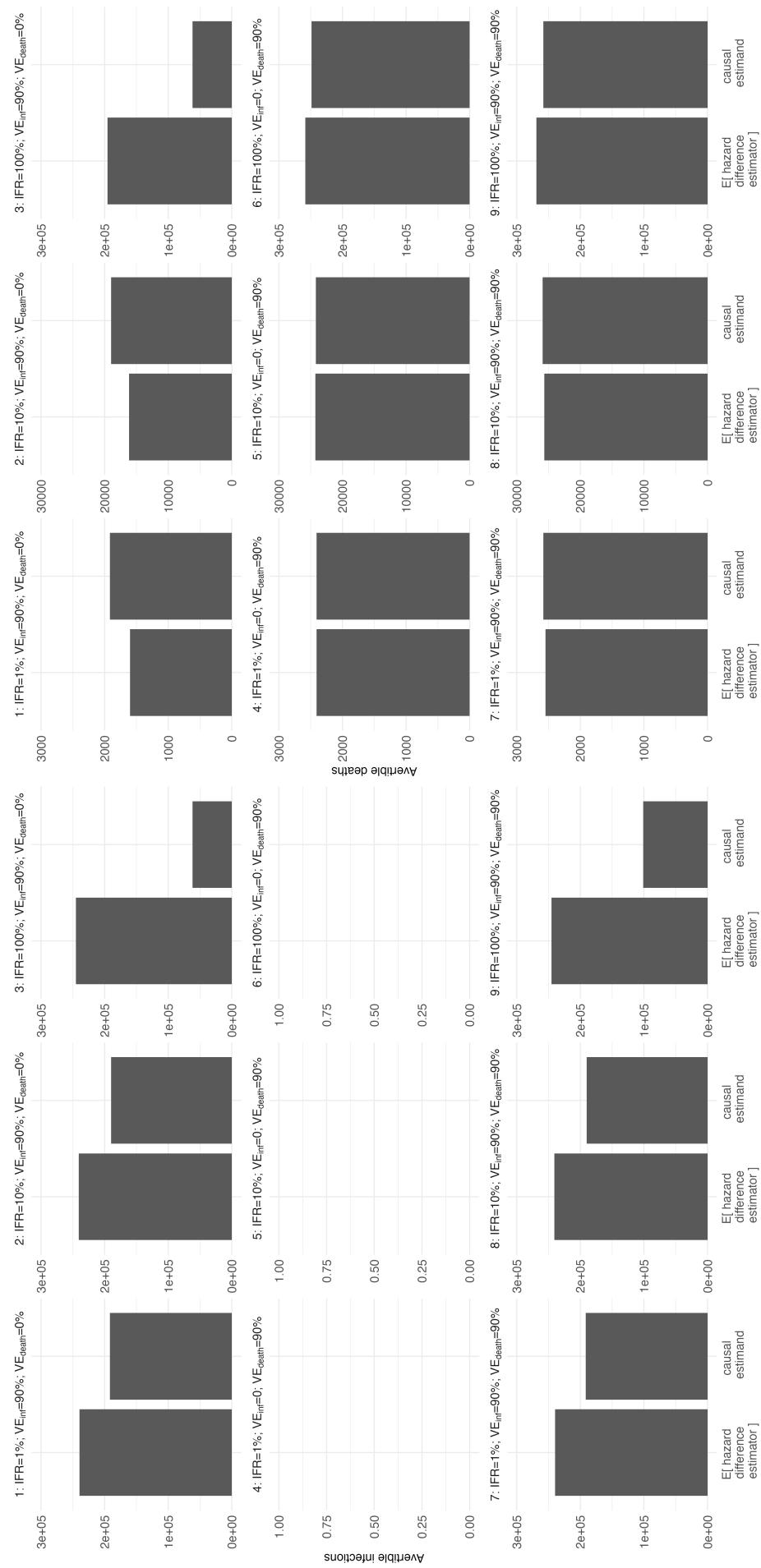
	<b>Scenario 1</b>	<b>Scenario 2</b>	<b>Scenario 3</b>	<b>Scenario 4</b>	<b>Scenario 5</b>	<b>Scenario 6</b>	<b>Scenario 7</b>	<b>Scenario 8</b>	<b>Scenario 9</b>
IFR: 1%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 0%	IFR: 10%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 0%	IFR: 100%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 0%	IFR: 10%, VE <sub>inf</sub> = 0%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 0%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 0%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 90%	IFR: 100%, VE <sub>inf</sub> = 90%, VE <sub>death</sub> = 90%
Avertible infections <sup>a</sup>	48041.54 (25.08%)	51050.02 (26.9%)	183221.65 (295.78%)	0.00	0.00	0.00	48034.29 (25.07%)	50967.59 (26.85%)	144266.72 (142.84%)
Avertible deaths <sup>a</sup>	-316.36 (-16.51%)	-2814.54 (-14.83%)	133578.07 (215.64%)	0.7 (0.03%)	72.59 (0.3%)	9981.36 (4.02%)	-31.53 (-1.22%)	-270.1 (-1.04%)	10765.53 (4.17%)

<sup>a</sup> Percentage differences were calculated using values rounded to nine decimal places, to avoid uninterpretable, extremely small non-zero values (absolute value  $< 10^{-9}$ ). All values in the table are rounded to two decimal places

## Infection

B

## Death



**FIGURE S7 |** Infections (A) and deaths (B) directly avertible among unvaccinated individuals with strategy  $\mathbf{p}'' = \{99, 7, 0, 01\}$  under different scenarios varied by infection-fatality rate (IFR), vaccine efficacy against infection ( $VE_{inf}$ ) and vaccine efficacy against death given infection ( $VE_{death}$ )

## References

1. Haas EJ, McLaughlin JM, Khan F, et al. Infections, hospitalisations, and deaths averted via a nationwide vaccination campaign using the Pfizer–BioNTech BNT162b2 mRNA COVID-19 vaccine in Israel: a retrospective surveillance study. *The Lancet Infectious Diseases*. 2022;22(3):357-366. doi:10.1016/S1473-3099(21)00566-1
2. Brault A, Hart A, Uribe P, et al. Direct impact of COVID-19 vaccination in Chile: averted cases, hospitalizations, ICU admissions, and deaths. *BMC Infect Dis*. 2024;24(1):467. doi:10.1186/s12879-024-09304-1
3. Santos CVBD, Noronha TGD, Werneck GL, Struchiner CJ, Villela DAM. Estimated COVID-19 severe cases and deaths averted in the first year of the vaccination campaign in Brazil: a retrospective observational study. *The Lancet Regional Health - Americas*. 2023;17:100418. doi:10.1016/j.lana.2022.100418
4. Kayano T, Sasanami M, Kobayashi T, et al. Number of averted COVID-19 cases and deaths attributable to reduced risk in vaccinated individuals in Japan. *The Lancet Regional Health - Western Pacific*. 2022;28:100571. doi:10.1016/j.lanwpc.2022.100571
5. Kayano T, Nishiura H. Assessing the COVID-19 vaccination program during the Omicron variant (B.1.1.529) epidemic in early 2022, Tokyo. *BMC Infect Dis*. 2023;23(1):748. doi:10.1186/s12879-023-08748-1
6. Vilches TN, Moghadas SM, Sah P, et al. Estimating COVID-19 Infections, Hospitalizations, and Deaths Following the US Vaccination Campaigns During the Pandemic. *JAMA Netw Open*. 2022;5(1):e2142725. doi:10.1001/jamanetworkopen.2021.42725
7. Schneider EC, Shah A, Sah P, et al. Impact of U.S. COVID-19 Vaccination Efforts: An Update on Averted Deaths, Hospitalizations, and Health Care Costs Through March 2022. To the Point. April 8, 2022. Accessed December 10, 2023. <https://doi.org/10.26099/d3dm-fa91>
8. Gavish N, Yaari R, Huppert A, Katriel G. Population-level implications of the Israeli booster campaign to curtail COVID-19 resurgence. *Sci Transl Med*. 2022;14(647):eabn9836. doi:10.1126/scitranslmed.abn9836
9. Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*. 2022;22(9):1293-1302. doi:10.1016/S1473-3099(22)00320-6
10. Shoukat A, Vilches TN, Moghadas SM, et al. Lives saved and hospitalizations averted by COVID-19 vaccination in New York City: a modeling study. *The Lancet Regional Health - Americas*. 2022;5:100085. doi:10.1016/j.lana.2021.100085
11. Hudgens MG, Halloran ME. Toward Causal Inference With Interference. *Journal of the American Statistical Association*. 2008;103(482):832-842. doi:10.1198/016214508000000292

12. Sobel ME. What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference. *Journal of the American Statistical Association*. 2006;101(476):1398-1407. doi:10.1198/016214506000000636
13. VanderWeele TJ, Tchetgen Tchetgen EJ. Effect partitioning under interference in two-stage randomized vaccine trials. *Statistics & Probability Letters*. 2011;81(7):861-869. doi:10.1016/j.spl.2011.02.019
14. Jia KM, Boyer CB, Wallinga J, Lipsitch M. Causal Estimands for Analyses of Averted and Avertible Outcomes due to Infectious Disease Interventions. *Epidemiology*. 2025;36(3):363-373. doi:10.1097/EDE.0000000000001839
15. Forastiere L, Airoldi EM, Mealli F. Identification and Estimation of Treatment and Interference Effects in Observational Studies on Networks. *Journal of the American Statistical Association*. 2021;116(534):901-918. doi:10.1080/01621459.2020.1768100
16. Yang W, Lipsitch M, Shaman J. Inference of seasonal and pandemic influenza transmission dynamics. *Proc Natl Acad Sci USA*. 2015;112(9):2723-2728. doi:10.1073/pnas.1415012112
17. Guerra FM, Bolotin S, Lim G, et al. The basic reproduction number (R<sub>0</sub>) of measles: a systematic review. *The Lancet Infectious Diseases*. 2017;17(12):e420-e428. doi:10.1016/S1473-3099(17)30307-9
18. Locatelli I, Trächsel B, Rousson V. Estimating the basic reproduction number for COVID-19 in Western Europe. Khudyakov YE, ed. *PLoS ONE*. 2021;16(3):e0248731. doi:10.1371/journal.pone.0248731
19. Nguyen VH, Crépey P, Williams BA, et al. Modeling the impact of early vaccination in an influenza pandemic in the United States. *npj Vaccines*. 2025;10(1):62. doi:10.1038/s41541-025-01081-5
20. Portnoy A, Jit M, Ferrari M, Hanson M, Brenzel L, Verguet S. Estimates of case-fatality ratios of measles in low-income and middle-income countries: a systematic review and modelling analysis. *The Lancet Global Health*. 2019;7(4):e472-e481. doi:10.1016/S2214-109X(18)30537-0
21. Ayoub HH, Mumtaz GR, Seedat S, Makhoul M, Chemaitelly H, Abu-Raddad LJ. Estimates of global SARS-CoV-2 infection exposure, infection morbidity, and infection mortality rates in 2020. *Global Epidemiology*. 2021;3:100068. doi:10.1016/j.gloepi.2021.100068
22. Centers for Disease Control and Prevention. *U.S. Flu Vaccine Effectiveness (VE) Data for 2021-2022.*; 2024. Accessed July 3, 2025. <https://www.cdc.gov/flu-vaccines-work/php/effectiveness-studies/2021-2022.html>
23. Di Pietrantonj C, Rivetti A, Marchione P, Debalini MG, Demicheli V. Vaccines for measles, mumps, rubella, and varicella in children. Cochrane Acute Respiratory Infections Group, ed. *Cochrane Database of Systematic Reviews*. 2021;2021(11). doi:10.1002/14651858.CD004407.pub5

24. Polack FP, Thomas SJ, Kitchin N, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med.* 2020;383(27):2603-2615. doi:10.1056/NEJMoa2034577
25. Ferdinand JM, Thompson MG, Blanton L, Spencer S, Grant L, Fry AM. Does influenza vaccination attenuate the severity of breakthrough infections? A narrative review and recommendations for further research. *Vaccine.* 2021;39(28):3678-3695. doi:10.1016/j.vaccine.2021.05.011
26. Nic Lothain LM, De Gier B, Van Der Maas N, et al. Immunogenicity, effectiveness, and safety of measles vaccination in infants younger than 9 months: a systematic review and meta-analysis. *The Lancet Infectious Diseases.* 2019;19(11):1235-1245. doi:10.1016/S1473-3099(19)30395-0
27. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 Vaccine Protection against SARS-CoV-2 Infection in Qatar. *N Engl J Med.* 2021;385(24). doi:10.1056/NEJMoa2114114
28. Cauchemez S, Ferguson NM. Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *J R Soc Interface.* 2008;5(25):885-897. doi:10.1098/rsif.2007.1292
29. Della Rossa F, Salzano D, Di Meglio A, et al. A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nat Commun.* 2020;11(1):5106. doi:10.1038/s41467-020-18827-5
30. Centers for Disease Control and Prevention. Rates of COVID-19 Cases or Deaths by Age Group and Vaccination Status. Published online 2025. Accessed August 14, 2025. [https://data.cdc.gov/Public-Health-Surveillance/Rates-of-COVID-19-Cases-or-Deaths-by-Age-Group-and/3rge-nu2a/about\\_data](https://data.cdc.gov/Public-Health-Surveillance/Rates-of-COVID-19-Cases-or-Deaths-by-Age-Group-and/3rge-nu2a/about_data)
31. Hernán MA, Sterne JAC, Higgins JPT, Shrier I, Hernández-Díaz S. A Structural Description of Biases That Generate Immortal Time. *Epidemiology.* 2025;36(1):107-114. doi:10.1097/EDE.0000000000001808
32. Jia KM, Hanage WP, Lipsitch M, et al. Estimated preventable COVID-19-associated deaths due to non-vaccination in the United States. *Eur J Epidemiol.* Published online April 24, 2023. doi:10.1007/s10654-023-01006-3
33. Zhong M, Glazer T, Kshirsagar M, et al. Estimating Vaccine-Preventable COVID-19 Deaths Among Adults Under Counterfactual Vaccination Scenarios in The United States: A Modeling Study Using Observational Data. *J Pharm Pharmacol Res.* 2023;07(03). doi:10.26502/fjppr.079