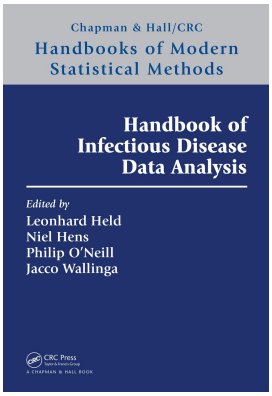

Handbook of Infectious Disease Data Analysis

Leonhard Held, Niel Hens, Philip O'Neill, Jacco Wallinga

Analysis of Vaccine Studies and Causal Inference

Publication details
https://www.routledgehandbooks.com/doi/10.1201/9781315222912-8
M. Elizabeth Halloran
Published online on: 04 Nov 2019

8

Analysis of Vaccine Studies and Causal Inference

M. Elizabeth Halloran

CONTENTS

8.1 Introduction

8.1.1 Background

The analysis of vaccine studies enjoys a long history. In 1915, Greenwood and Yule (1915) published an 85-page treatise on "The Statistics of Anti-typhoid and Anti-cholera Inoculations, and the Interpretation of such Statistics in general." In the days before randomization, they stated three conditions for valid inference when estimating protective effects of vaccination (inoculation). "1. The persons must be, *in all material respects*, alike. 2. The effective

exposure to the disease must be identical in the case of inoculated and uninoculated persons. 3. The criteria of the fact of inoculation and of the fact of the disease having occurred must be independent." Since the advent of randomized trials, randomization is supposed to ensure the validity of the comparison groups. Greenwood and Yule also discussed heterogeneity in susceptibility and protection and the role of a possible immune threshold level for protection, issues that are still discussed today.

Historically, the focus was on estimating the protective effects of vaccines in vaccinated individuals. In 1916, Sir Ronald Ross (1916) published his general treatise on the Theory of Happenings. He separated different kinds of happenings into two classes, namely "(a) those in which the frequency of the happening is *independent* of the number already affected; and (b) those in which the frequency of the happening depends on this quantity." Infectious diseases belong to the latter. Due to dependent happenings, in infectious diseases, interventions, including vaccination, can have several kinds of effects in populations, not just protective effects in individuals, which require different study designs and methods of analysis for their evaluation.

Halloran and Struchiner (1991, 1995) defined study designs for dependent happenings that allow evaluation of the direct, indirect, total, and overall effects of vaccination, and other infectious disease interventions (Figure 8.1). Consider two clusters, or populations, of individuals. In one of the populations, a certain portion of individuals is vaccinated and the rest remain unvaccinated. In the other population, no one is vaccinated. The *direct effect* of vaccination in the population in which some individuals were vaccinated is defined by comparing the average outcomes in vaccinated individuals with the average outcomes in unvaccinated individuals. The *indirect effects*, also known as *spillover effects*, are defined as a contrast between the average outcomes in unvaccinated individuals in the population

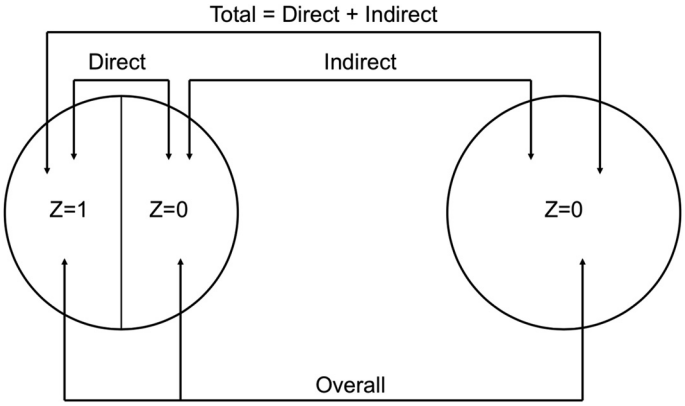

FIGURE 8.1
Study designs for dependent happenings. Two clusters, or populations, are considered under two different scenarios. In the scenario on the left, a certain portion of individuals in the cluster receive vaccination (or other treatment), $Z = 1$, and the other portion of individuals receive control, $Z = 0$. In the scenario on the right, everyone receives control. Control could be current best practice, a placebo, or nothing. The direct, indirect, total, and overall effects of intervention are defined by the indicated contrasts. The figure makes no reference to an assignment mechanism, comparability of possible subgroups within the populations, or inference about the different effects. (Adapted from Halloran, M.E. and Struchiner, C.J. *Epidemiology*, 2, 331–338, 1991; Halloran, M.E. and Struchiner, C.J. *Epidemiology*, 6, 142–151, 1995.)

with vaccination and the average outcomes of unvaccinated individuals in the unvaccinated population. The *total effects* are defined by comparing the average outcomes in the vaccinated individuals in the vaccinated population to the average outcomes in the unvaccinated individuals in the unvaccinated population. The *overall effects* are defined by the contrast in the average outcomes in the entire population where some individuals were vaccinated compared to the average outcomes of the entire population that did not receive vaccine.

8.1.2 Causal inference

In recent decades, causal inference methods have been developed for evaluating some effects of vaccination. Causal inference based on potential outcomes (Rubin, 1980, 1986; Holland, 1986) is a principled framework for carefully defining the causal effects of interventions. A potential outcome is defined as the outcome an individual would have if the individual were to receive a particular intervention. The potential outcomes are generally assumed fixed prior to an individual actually receiving treatment. That is why they are called potential, or counterfactual, outcomes. A common assumption in causal inference is that the treatment assignment in one individual does not affect the potential outcomes in another individual, called the assumption of no interference (Cox, 1958). The Stable Unit Treatment Value Assumption (SUTVA) contains the assumption of no interference, as well as the assumption that all treatments and their potential outcomes are represented (Rubin, 1980). Assuming no interference, if there are only two treatments, such as a vaccine and a control, then an individual has only two potential outcomes.

Causal effects can be defined at the individual level as a contrast of potential outcomes under two different interventions. For example, the individual causal effect of treatment compared to a control is the contrast between the individuals's potential outcome under treatment and the individual's potential outcome under control. Generally only one potential outcome can be observed in an individual, so the individual causal effect cannot be observed. The population average causal effect is defined as the contrast of the average of the difference between outcomes if everyone received treatment and if everyone received control. This measure also is not possible to observe.

Drawing inference about treatment effects requires specification of an assignment mechanism of the treatments to individuals or knowledge of how individuals select their treatment. An assignment mechanism where the treatment assigned is independent of the potential outcomes, such as randomization, is a useful assignment. Under the assumption of no interference between individuals and randomization, we can construct an unbiased estimator of the population average causal effect from the contrast of the observed average outcomes in the two treatment groups. The assumption of randomization to specify estimators of the estimands of interest is our point of departure for estimating effects of interest. Observational studies in which vaccine is not assigned randomly may be subject to biases, but can be viewed as departures from randomized experiments.

Obviously, in interventions for infectious diseases, such as vaccination, the intervention assignment of one individual might affect the potential outcomes of another individual. For example, if an individual is vaccinated, the individual may not become infected or be less infectious if infected, and then not infect another individual that he or she would have infected had he or she not been vaccinated. Thus, the assumption of no interference may be violated, which we consider further herein.

Causal inference methods assuming no interference are useful in understanding vaccine effects on post-infection outcomes (Section 8.2.6) and evaluating surrogates of vaccine-induced protection (Section 8.4). Causal inference where interference may be present is useful in defining estimands and estimators of different population-level effects of vaccination

as in Figure 8.1 (Section 8.3.2). Section 8.3.3.1 presents a case study of evaluating and interpreting such estimates. Evaluating the effects of vaccination on infectiousness requires a combination of causal methods (Section 8.2.7).

8.1.3 Focus of this chapter

In this chapter, the focus is on evaluating prophylactic vaccines in populations. However, many of the principles outlined in this chapter could apply to other infectious disease interventions, either prophylactic or treatment, or pharmaceutical or behavioral, or combinations of interventions. The book *Design and Analysis of Vaccine Studies* by Halloran et al. (2010) covers some of the material presented here in much more detail. The book *Cluster Randomised Trials* by Hayes and Moulton (2017, second edition) is an excellent resource for the practical design and analysis of cluster-randomized studies, particularly in the infectious disease context. In this chapter, we present some key ideas and recent topics in vaccine studies and causal inference for evaluating vaccination.

8.2 Estimating Different Effects of Vaccination

We present a framework showing the relation among many of the different vaccination effects and the study designs to estimate them. Commonly, vaccine efficacy and effectiveness (VE) are estimated as 1 minus some measure of relative risk, RR:

$$VE = 1 - RR.$$

However, vaccine effects also can be measured on the risk difference scale, the odds ratio scale, and relative risk scale. The vaccine effect of interest determines the choice of comparison groups, the unit of observation, the choice of parameter, and the amount of information about the transmission system required for estimation. The framework draws on the dependent happening relation in infectious diseases. Table 8.1 list several different vaccine effects. Table 8.2 provides a framework for understanding how many of the effects and estimators of different vaccine effects relate to one another.

TABLE 8.1
Some vaccine effects of interest

Symbol	Definition
VE_S	Vaccine efficacy for susceptibility
VE_{SP}	Vaccine efficacy for susceptibility to disease
VE_{col},VE_{acq}	Vaccine efficacy for colonization (acquisition)
VE_P	Vaccine efficacy for progression or pathogenicity
VE_I	Vaccine efficacy for infectiousness
VE_T	Total vaccine efficacy
$\text{VE}_{Indirect}$	Indirect effects of vaccination in those not vaccinated
VE_{Total}	Total effects of vaccination in those vaccinated
$\text{VE}_{Overall}$	Overall population-level effects

8.2.1 Protective effects of vaccination, VE_S, VE_{SP}

VE_S, the vaccine efficacy for susceptibility, is a measure of how protective vaccination is against infection. VE_{SP} denotes vaccine efficacy against disease. With most infectious agents, the major interest has been in preventing clinical illness. In studies of vaccines against such infectious agents, including measles, influenza, dengue, and pertussis, ascertainment of the outcome of interest is often on individuals who have symptomatic disease, who then might have a laboratory test done to confirm the infectious agent under study. In this scenario, asymptomatic infections would not be ascertained.

VE_P, vaccine efficacy for progression or pathogenicity, measures the efficacy of vaccination in preventing a post-infection outcome. Depending on the situation, the measure of interest can be the effect of prophylactic vaccination on the rate or probability of progressing to disease, conditional on being infected. If ascertainment is on disease, VE_P could be a measure of the effect of vaccination on the probability of severe disease. Although VE_S, VE_{SP}, and VE_P are all measures of the direct protective effects of vaccination, there is an important difference. Studies to estimate VE_S and VE_{SP} are based on an outcome in participants who are susceptible to infection, whereas studies to estimate VE_P are based on an outcome in participants who are already infected. The denominators in the two different types of studies are different. In randomized studies, as long as the outcome is the first outcome of interest after randomization, whether infection, VE_S, or disease, VE_{SP}, the validity of the comparison is preserved. This result is not necessarily true with VE_P (Section 8.2.6).

We first consider study designs for estimating the protective effects of vaccination, VE_S (VE_{SP}). In the following, we use VE_S to denote both VE_S and VE_{SP}, that is, it could be ascertainment on infection or disease, unless there is a need to differentiate the two. In Table 8.2, these are represented in the column labeled "Susceptibility." The estimates of VE_S are obtained from the relative risk of infection or disease in the vaccinated individuals compared to the unvaccinated individuals:

$$\mathrm{VE}_S = 1 - \frac{R(\text{vaccinated people})}{R(\text{unvaccinated people})},$$

where R denotes one of the measures of risk. The measure of risk can be a form of the transmission probability, which conditions on exposure to infection, or the incidence rate, hazard rate, or cumulative incidence (attack rate), which do not condition on exposure to infection. In Table 8.2, the amount of information about the transmission system required for the efficacy estimates decreases from Level I in the top row to Level IV in the bottom row.

The top row of Table 8.2 contains measures of VE that rely on information about exposure to infection and contacts between infectious individuals and susceptible individuals. The first is a measure of VE_S based on the transmission probability, $\mathrm{VE}_{S,p}$. Let the *transmission probability*, denoted p_{ij}, be the probability that, conditional upon a contact between an infective source with covariate status i and a susceptible host with covariate status j, successful transfer and establishment of the infectious agent will occur. A related concept is the *secondary attack rate* (SAR_{ij}), defined as the proportion of susceptible individuals with covariate status j making contact with an infectious person of covariate status i that becomes infected. The SAR is a special case of the transmission probability.

Let 0 and 1 denote being unvaccinated and vaccinated. Then, for example, p_{01} denotes the transmission probability per contact from an unvaccinated infective person to a vaccinated uninfected person. Let $p_{\cdot 0}$ and $p_{\cdot 1}$ denote the transmission probability to unvaccinated and vaccinated susceptible individuals, where the dot in the subscript can denote any vaccine status or an average across the population.

TABLE 8.2

Parameters used for measuring various effects of vaccination. The levels form a hierarchy, with higher levels requiring less information about the transmission system, with only level I requiring actual contact information[a]

		Comparison Groups and Effect			
Level	**Parameter Choice**	**Susceptibility**	**Infectiousness**	**Combined Change in Susceptibility and Infectiousness**	
	Conditional on exposure to infection:				
I	Transmission probability, p Secondary attack rate, SAR	$\mathrm{VE}^{\mathrm{b}}_{S,p} = 1 - \frac{p_{\cdot 0}}{p_{\cdot 1}}$	$\mathrm{VE}_{I,p} = 1 - \frac{p_{1\cdot}}{p_{0\cdot}}$	$\mathrm{VE}_{T,p} = 1 - \frac{p_{11}}{p_{00}}$	
		Study Design			
		Direct	Indirect	Total	Overall
	Unconditional:				
II	Incidence rate, IR	$\mathrm{VE}_{S,IR} = 1 - \frac{\mathrm{IR}_{A1}}{\mathrm{IR}_{A0}}$	$\mathrm{VE}_{Indirect,IR} = 1 - \frac{\mathrm{IR}_{A0}}{\mathrm{IR}_{B0}}$	$\mathrm{VE}_{Total,IR} = 1 - \frac{\mathrm{IR}_{A1}}{\mathrm{IR}_{B0}}$	$\mathrm{VE}_{Overall,IR} = 1 - \frac{\mathrm{IR}_{A\cdot}}{\mathrm{IR}_{B\cdot}}$
	Hazard rate, λ	$\mathrm{VE}_{S,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{A0}}$	$\mathrm{VE}_{Indirect,\lambda} = 1 - \frac{\lambda_{A0}}{\lambda_{B0}}$	$\mathrm{VE}_{Total,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{B0}}$	$\mathrm{VE}_{Overall,\lambda} = 1 - \frac{\lambda_{A\cdot}}{\lambda_{B\cdot}}$
III	Proportional hazards, PH	$\mathrm{VE}_{S,PH} = 1 - e^{\beta_1}$	$\mathrm{VE}_{Indirect,PH} = 1 - e^{\beta_{ind}}$	$\mathrm{VE}_{Total,PH} = 1 - e^{\beta_{tot}}$	$\mathrm{VE}_{Overall,PH} = 1 - e^{\beta_{over}}$
IV	Cumulative incidence, CI Attack rates, AR	$\mathrm{VE}_{S,CI} = 1 - \frac{\mathrm{CI}_{A1}}{\mathrm{CI}_{A0}}$	$\mathrm{VE}_{Indirect,CI} = 1 - \frac{\mathrm{CI}_{A0}}{\mathrm{CI}_{B0}}$	$\mathrm{VE}_{Total,CI} = 1 - \frac{\mathrm{CI}_{A1}}{\mathrm{CI}_{B0}}$	$\mathrm{VE}_{Overall,CI} = 1 - \frac{\mathrm{CI}_{A\cdot}}{\mathrm{CI}_{B\cdot}}$

[a] Adapted from Halloran et al. (1997). The subscripts 0 and 1 denote unvaccinated and vaccinated people, respectively. Population A contains both vaccinated and unvaccinated people. All people in population B are unvaccinated (Figure 8.1, left = A, right = B). The subscripts S, I, and T denote susceptibility, infectiousness, and combined effects. The Cox proportional hazards estimator is denoted by e^{β} with the corresponding subscript. Time has been omitted from the table for notational clarity.

[b] Vaccine efficacy/effectiveness

Then $\text{VE}_{S,p}$ based on the transmission probability(or secondary attack rate) (Table 8.2, top row) is estimated from

$$VE_{S,p} = 1 - \frac{p_{.1}}{p_{.0}} = 1 - \frac{\frac{\text{vaccinated infections}}{\text{vaccinated contacts}}}{\frac{\text{unvaccinated infections}}{\text{unvaccinated contacts}}} \tag{8.1}$$

Estimating vaccine efficacy from the transmission probability ratios requires information on who is infectious and when, and whom they contact and how. The concept of a *contact* is very broad and must be defined in each particular study. Often it is defined for individuals within a small transmission unit such as a household or sexual partnership. Analysis of such studies can either assume that the transmission units are independent, such as in the conventional SAR analysis, or that the transmission units are embedded in a community where individuals are exposed to infection within and outside the transmission unit. The latter can be based on final value data or longitudinal data. These topics are handled elsewhere in this book and not considered further here. In the top row of Table 8.2, the vaccine efficacy for infectiousness, $\text{VE}_{I,p}$, is the relative reduction in the transmission probability from a vaccinated infected person to a susceptible person compared to the transmission probability from an unvaccinated infected person to a susceptible person. The $\text{VE}_{T,p}$ is the relative reduction in the transmission probability if both the infected person and the susceptible person in a contact are vaccinated compared to if both the infected person and the susceptible person in a contact are unvaccinated. VE_I and VE_T are discussed in Section 8.2.7.

Standard parameters for estimating VE_S that do not require exposure to infection information are incidence rates, hazard rates, or cumulative incidence (attack rate). Statistical analysis for estimates of these can be done using standard software, such as the R statistical programming system (R Core Team 2018). Primary vaccine efficacy studies often report $\text{VE}_{S,IR}$ based on the relative number of events per person-time,

$$\text{VE}_{S,IR} = 1 - \frac{\text{vaccinated events/person-time}}{\text{unvaccinated events/person-time}}. \tag{8.2}$$

The usual assumption is that the numbers of events follow a Poisson distribution. Similarly, investigators may estimate the hazard rates at time t in the vaccinated and unvaccinated, $\lambda_1(t)$ and $\lambda_0(t)$, using survival analysis methods. Then the VE_S is based on the hazard rate ratio

$$\text{VE}_{S,\lambda}(t) = 1 - \frac{\lambda_1(t)}{\lambda_0(t)}. \tag{8.3}$$

When covariates such as age and gender are added, the analyses are stratified by the covariates or Poisson regression can be used. Under the assumption that the vaccine effect is multiplicative, constant, and homogeneous, the Cox proportional hazards model can be used to estimate $\text{VE}_{S,PH}$. In this case, it is not necessary to estimate the hazard rate in the unvaccinated group, but only the relative hazard rate. This situation requires only the ordering of the infection times. Covariates, including time-dependent covariates, can easily be incorporated using standard software, such as the `survival` package in R (Therneau, 2015).

$\text{VE}_{S,CI}(T)$ based on the cumulative incidence uses only information about whether persons are infected by the end of the study at time T, that is, final value data:

$$\begin{aligned} VE_{S,CI}(T) &= 1 - \frac{\text{vaccinated infection events/persons–at–risk}}{\text{unvaccinated infection events/persons–at–risk}} \\ &= 1 - \frac{CI_1(T)}{CI_0(T)}\ . \end{aligned} \tag{8.4}$$

Using the simple $\text{VE}_{S,CI}(T)$ in Equation (8.4) assumes there is no loss to follow-up. However, we also can estimate the cumulative incidence from time-to-event data that allows for censoring. This method may be preferable to using the Poisson model in Equation (8.2) or the Cox regression model because those are both based on parametric assumptions that may fail. Vaccine efficacy at time t based on the cumulative incidence function is

$$VE_{S,CI}(t) = 1 - \frac{CI_1(t)}{CI_0(t)} \ . \tag{8.5}$$

The estimated vaccine efficacy with simultaneous confidence bands can be plotted as a function of time to visualize the behavior of vaccine efficacy over time.

Neafsey et al. (2015) presented estimates of vaccine efficacy using both 1 minus the ratio (vaccine vs. control) of cumulative incidences and 1 minus the ratio (vaccine vs. control) of hazards, the latter under the assumption that the hazards are proportional over time. The context was to estimate the efficacy of the RTS,S/AS01 malaria vaccine according to genetic diversity of the parasite. The estimates of vaccine efficacy based on the two methods differ significantly in several of the analyses.

If there is no adjustment for covariates, the cumulative incidence function can be estimated by the nonparametric likelihood estimator (Aalen, 1978), implemented in the R `cmprsk` package, based on the Kaplan-Meier estimate of the survival function and the Nelson-Aalen estimates of the cumulative hazard functions. Adjusted analyses that include covariates to estimate the cumulative incidence functions can use targeted maximum likelihood estimation that leverages information in covariates to improve efficiency (Benkeser et al. 2018).

8.2.2 Hierarchy of VE_S parameters

The different VE_S parameters require different levels of information for their estimation and make different demands on study design and data collection (Rhodes et al., 1996) [65]. Levels I through IV in Table 8.2 form a hierarchy, with higher levels requiring less information about the transmission system, and only Level I requiring contact and exposure to infection information. Because of the dependent happening structure of events in infectious diseases, an intrinsic relation exists among the different parameters on which the VE_S estimators are based.

Let p_{ij} be the transmission probability as previously defined. Let c denote the contact rate in a population assuming that people are randomly mixing, and let $P(t)$ denote the prevalence of infectives at time t. Then the hazard rate $\lambda(t)$ (or incidence rate or force of infection) at time t can be expressed as the product of the contact rate, the transmission probability, and the probability that a contact is infectious:

$$\lambda(t) = cp_{ij}P(t). \tag{8.6}$$

So even if the different components of the hazard rate are not measured, we can consider the underlying process that is producing the infections we observe. Similarly, the cumulative incidence, $CI(T)$, at some time T is a function of the hazard rate during the follow-up period, and thus also a function of the transmission probability, contact rate, and prevalence of infection in the contacts.

8.2.3 Vaccine efficacy for colonization and acquisition

Many infectious agents, such as *Streptococcus pneumoniae* (pneumococcus), meningococcus, and *Hemophilus influenzae b* bacteria, colonize the nose and throat passages without causing

overt disease. The state of being colonized is also called nasopharyngeal carriage. Colonized individuals generally are asymptomatic, but they play a central role in transmission. They can transmit to other susceptible individuals who in their turn may develop severe disease after being colonized. Individuals can clear the bacteria and then be re-colonized repeatedly. Pneumococcus has over 90 strains. A common pneumococcus vaccine contains antigens to 13 strains, although one is available with even more. $\mathrm{VE}_{\mathrm{col}}$ measures the efficacy against colonization (Auranen et al., 2000, Käyhty et al., 2006). In particular, the rate of acquisition of pneumococcal carriage also can be used as the outcome measure to estimate vaccine efficacy, $\mathrm{VE}_{\mathrm{acq}}$ (Rinta-Kokko et al., 2009).

Because carriage is asymptomatic, studies need to do active swabbing of individuals to determine if they are colonized or not. These studies are generally longitudinal, expensive, and invasive, thus interest has developed to estimate $\mathrm{VE}_{\mathrm{acq}}$ from just one observation of current status data. Auranen et al. (2013) present two examples of vaccine studies estimating vaccine efficacy from just cross-sectional data on nasopharyngeal colonization by pneumococcus. They present a framework for defining and estimating strain-specific and overall vaccine efficacy for susceptibility to acquisition of colonization ($\mathrm{VE}_{\mathrm{acq}}$) when there is a large number of strains with mutual interactions and recurrent dynamics of colonization, such as in pneumococcus. They develop estimators based on one observation of the current status per study subject, evaluate their robustness, and re-analyze the two vaccine trials.

Mehtälä et al. (2016) consider the problem of estimating heterogeneous vaccine efficacy against an infection that can be acquired multiple times. Estimation is based on a limited number of repeated measurements of the current status of each individual. They investigate how the choice of time intervals between consecutive samples affects the estimability and efficiency of vaccine efficacy parameters. The authors suggest practical guidelines that allow estimation of all components. For situations in which the estimability of individual components fails, they suggest using summary measures of vaccine efficacy.

8.2.4 Modes of action and time-varying VE_S

Smith et al. (1984) pointed out that the choice of method of analysis may depend on the mode of vaccine protection. They defined Type I and Type II modes of action. In the Type I mechanism, vaccination is assumed to reduce the instantaneous disease rate in all the vaccinated people by a constant proportion. That is, Type I has a multiplicative effect on the baseline hazard. In the Type II mechanism, a certain proportion of the vaccinated individuals are completely protected while the other individuals have no protection, and essentially are at risk like the unvaccinated individuals. Struchiner et al. (1990) used the term "leaky" from the malaria vaccinology literature to replace the term Type I and Halloran et al. (1991) used the term "all-or-none" to replace the term Type II. The assumptions about the mechanism of vaccine protection have implications for the choice of efficacy measures. Assuming the vaccine effect does not wane, in general, estimates of VE_{CI} will be time-invariant under the all-or-none (Type II) mechanism, and estimates of VE_{IR} or VE based on the Cox proportional hazards model will be time-invariant under a leaky (Type I) mechanism. Frailty models can be used for analysis if a combination of leaky and all-or-none is suspected (Halloran et al., 1996).

To estimate waning vaccine efficacy, Durham et al. (1998) used a method based on the smoothed scaled Schoenfeld residuals (Grambsch and Therneau, 1994). Cox models can be fitted using the `survival` package (Therneau, 2015) in the R statistical programming system (R Core Team 2018). Plots of the time trends of vaccine efficacy can be made using the R function VEplot in the `kyotil` package (Fong and Sebestyen, 2018).

The approach in Equation (8.5) to estimate and graph $\text{VE}_{S,CI}(t)$ allows us to visually see how vaccine efficacy changes over time, as in Neafsey et al. (2015). However, if the vaccine has a leaky (Type I) mechanism, the estimate of $\text{VE}_{S,CI}(t)$ based on the cumulative incidence function may change over time even though the vaccine efficacy does not wane.

8.2.5 Observational studies

Much has been written on the use of case-control studies in evaluating vaccine efficacy and effectiveness (Struchiner et al., 1990; Rodrigues and Smith, 1999). We do not cover that here. We highlight two recent developments for observational studies, namely, the test negative design and quasi-experimental designs.

8.2.5.1 Test negative design

The test-negative design to estimate VE_S enrolls cases presenting to a medical facility that have symptoms consistent with the disease of interest. Then a laboratory test is performed to confirm whether the case has the disease of interest or not. Those individuals who test positive are considered the cases and those who test negative are the non-cases. Vaccine efficacy is estimated from the odds ratio comparing the odds of testing positive for the disease among vaccinated individuals with the odds among unvaccinated individuals using logistic regression, adjusting for potential confounders. The method is fairly inexpensive if the tests are being performed anyway. The method has been used extensively for estimating the efficacy of influenza vaccines (Jackson and Nelson, 2013) and, more recently, of rotavirus vaccines in low resource settings (see, for example, Bar-Zeev et al., 2015). A core assumption of the method is that the vaccine does not have an effect on the causes of the symptomatic cases who test negative.

One method of examining how well the test negative design performs has been to use randomized, placebo controlled vaccine trials as a gold standard and compare the estimates that a test negative design would have yielded if it had been performed instead (De Serres et al., 2013). For example, Schwartz et al. (2017) derived test-negative vaccine effectiveness estimates from three large randomized placebo controlled trials of rotavirus vaccines in sub-Saharan Africa and Asia. They found that the test-negative design estimates of vaccine efficacy were nearly equivalent to those from the randomized controlled trials.

Although this result is comforting, it is not a true examination of the validity of the approach. The first use of causal inference to examine the theoretical basis of the test-negative study was by Sullivan et al. (2016). The test-negative design is generally considered to control for care-seeking behavior by using only cases that seek medical care, in contrast to case-control studies that use population-based controls (Jackson and Nelson, 2013). It also is assumed to reduce misclassification of cases by requiring laboratory confirmation of the infectious agent of interest. Using directed acyclic graphs, Sullivan et al. (2016) showed how bias may be reduced in some instances, but may be introduced in others. Directed acyclic graphs are one approach to causal inference. Richardson and Robins (2013) provided a detailed analysis of how causal inference with potential outcomes relates to that using directed acyclic graphs and developed a unified theory of single world intervention graphs (SWIGs) that lies outside the scope of this chapter. Westreich and Hudgens (2016) pointed out further that because the test-negative design is limited to those seeking health care, the study may not be generalizable to the population if vaccine efficacy differs by health-care seeking behavior or associated factors. In addition, the odds ratio is not collapsible in that the conditional causal odds ratio will not in general equal the marginal causal odds ratio (Greenland et al., 1999). Thus, further research is left to be done on the use and validity of the test-negative design for evaluating vaccines.

8.2.5.2 Quasi-experimental design

An approach as yet underutilized in public health (Bor et al., 2014; Moscoe et al., 2015) is the quasi-experimental design known as the regression discontinuity design (Thistlewaite and Campbell, 1960). In particular, the design might be useful for evaluating vaccines in some situations when randomized studies are not feasible or ethical. The regression discontinuity design allows for causal inference about the effects of interventions when certain conditions hold. The main idea is that an intervention, such as a vaccine, would be administered to a group based on an arbitrary continuous cut-off, such as age. Then we might assume that those just below the cutoff who do not receive vaccine would be comparable to those just above the cutoff who do receive vaccine. The approach has been used widely in economics studies where randomization is often unfeasible. It seems not yet to have been developed in the context of evaluating vaccines. Aronow et al (2016) consider the regression discontinuity design under interference using a local randomization approach, where the causal estimands are related to those presented in Section 8.3.

8.2.6 Vaccine effects on post-infection outcomes, VE_P

Sometimes interest is in the effect of a vaccine on an outcome that occurs after an individual is infected. For example, we might be interested in estimating the probability of developing a symptomatic case in vaccinated infected individuals compared to unvaccinated infected individuals. Then VE_P is defined as 1 minus the ratio of a summary measure of the post-infection outcome in the infected vaccinated individuals and a summary measure of the post-infection outcome in the infected unvaccinated individuals:

$$\mathrm{VE}_P = 1 - \frac{\dfrac{\text{vaccinated post-infection outcome}}{\text{infected vaccinated people}}}{\dfrac{\text{unvaccinated post-infection outcome}}{\text{infected unvaccinated people}}}. \tag{8.7}$$

Similarly, if a post-clinical outcome in the clinical cases is of interest, then VE_P is defined as 1 minus the ratio of a summary measure of the post-clinical outcome in the vaccinated cases and a summary measure of the post-clinical outcome in the unvaccinated cases:

$$\mathrm{VE}_P = 1 - \frac{\dfrac{\text{vaccinated post-clinical outcome}}{\text{vaccinated clinical cases}}}{\dfrac{\text{unvaccinated post-clinical outcome}}{\text{unvaccinated clinical cases}}}. \tag{8.8}$$

For example, we might be interested to estimate if the severity of disease is lower in vaccinated clinical cases than in unvaccinated clinical cases. Use of Equations (8.7) and (8.8) without further adjustment assumes that those who get infected in the vaccinated and unvaccinated groups are comparable.

However, conditioning on an event, such as infection, that occurs subsequent to receipt of vaccine or control, could result in selection bias, even under randomization. Issues related to interpreting malaria vaccine trials motivated Struchiner et al. (1994) to consider the problem of vaccinated and unvaccinated groups not being comparable after being infected even in randomized trials. With the development of human immunodeficiency virus (HIV) vaccine candidates, and HIV vaccine trials in which infection was actively ascertained, concern grew that the infected individuals in the vaccinated group and infected individuals in the unvaccinated group might not be comparable, leading to biased estimates of the effect of vaccination on post-infection outcomes (Hudgens et al., 2003; Gilbert et al., 2003).

For example, assume that the potential immune response to HIV infection has a distribution in the population before individuals are randomized to vaccine or control. Randomization would assure that in large samples, the potential distribution of the immune

response to HIV infection before vaccination would be the same in the vaccine and the control groups. However, vaccination might enhance protection only in individuals who have a stronger immune system, conferring some level of protection against infection if exposed. Then the individuals in the vaccinated group who become infected would be the ones with weaker immune systems, whereas the infected individuals in the unvaccinated group would be those with a weaker immune system as well as those with a stronger immune system. In this situation, if a post-infection outcome in the vaccinated group is compared with that in the unvaccinated group, the vaccine could appear to make things worse, even if vaccination has absolutely no effect on anything after infection.

For example, if individuals with a weaker immune system tend to have a higher viral load after being infected than those with a stronger immune system, then the mean viral load in the infected vaccinated group would be higher than the mean viral load in the infected unvaccinated group. This observation could lead to the false conclusion that the vaccine made the post-infection outcome worse, possibly resulting in rejection of a potentially useful vaccine candidate (Hudgens et al., 2003; Gilbert et al., 2003). However, the vaccine in this case actually does not make anything worse. The problem is that the infected vaccinated group and infected control group are no longer comparable because of selection bias.

Frangakis and Rubin (2002) proposed a method in the causal inference with potential outcomes framework to address this problem. The method, called *principal stratification*, adjusts for post-treatment variables by stratifying on the joint potential post-treatment variables under each of the treatments being considered. A principal stratum is composed of individuals with the same joint post-treatment variable. Causal effects are then defined within the principal strata. In fact, in our situation, the causal effect estimand for VE_P is only defined in the principal stratum in which individuals become infected regardless of being assigned to vaccine or control.

A difficulty in using this approach is that which individuals are in which principal stratum is generally not identifiable without further assumptions. Thus, methods have been developed to do sensitivity analyses and to establish bounds on the degree of selection bias that may be present. Several papers have been published using this approach to assess vaccine effects on post-infection outcomes. In studying HIV vaccines, Hudgens et al. (2003) and Gilbert et al. (2003) adopted the principal stratification approach to assess HIV vaccine effects on the continuous post-infection outcome viral load. Hudgens and Halloran (2006) developed methods for the causal vaccine effects on binary post-infection outcomes with applications to pertussis and rotavirus vaccines.

8.2.7 Vaccine efficacy for infectiousness, VE_I

A vaccinated individual who becomes infected might have a lower probability of transmitting to a susceptible individual during a contact than an unvaccinated individual who becomes infected, either because that individual is less infectious or infectious for a shorter period of time. The vaccine efficacy for infectiousness, VE_I, measures the relative reduction in the ability of a vaccinated infected individual compared to an unvaccinated infected individual to transmit the infectious agent to others. Estimating reduction in infectiousness can be of considerable public health interest, particularly with vaccines that do not protect well against infection.

In Table 8.2, VE_I is a measure that conditions on exposure to infection. When based on the transmission probability it can be estimated as

$$\mathrm{VE}_{I,p} = 1 - \frac{p_{1\cdot}}{p_{0\cdot}}. \tag{8.9}$$

The $\cdot$ indicates that the individual being exposed is not stratified by vaccine status. For example, in a study in Niakhar, Senegal, Préziosi and Halloran (2003) estimated the relative

reduction in infectiousness to household contacts of a vaccinated case of pertussis compared to an unvaccinated case, VE_I to be 67 percent (95% CI 29, 86). The analysis can be stratified by vaccine status of the exposed individual (Préziosi and Halloran, 2003). The combined effect of having both individuals in a contact being vaccinated compared to neither being vaccinated, denoted VE_T in Table 8.2, also can be estimated.

However, just as with VE_P in Section 8.2.6, VE_I compares the transmission probability only in infected vaccinated and infected unvaccinated individuals. These individuals might not be comparable for the same reasons even if the study is randomized, so estimates of VE_I can be subject to post-randomization selection bias (Hudgens and Halloran, 2006). Using the method described in Equation (8.9) without further adjustment assumes no selection bias. The situation is more complicated than with VE_P, because in estimating the transmission probability, the outcome is dependent on the infection status of another exposed individual.

VanderWeele and Tchetgen Tchetgen (2011), and Halloran and Hudgens (2012a) proposed causal quantities corresponding to the infectiousness effect in the simple situation of households of size two. Halloran and Hudgens (2012a) consider the general case that one or both individuals could be exposed outside the household as well as either, neither or both could be randomized to vaccine. The approach combines causal inference with interference (see Section 8.3.2) with principal stratification (Frangakis and Rubin, 2002). The latter accounts for the fact that the comparison in the groups who become infected may be subject to selection bias. The causal infectiousness effect is not identifiable without further assumptions, but bounds on the selection effects can be set (Halloran and Hudgens, 2012a, 2012b). VanderWeele et al. (2014) presented results for sensitivity analyses for causal infectiousness effects. These approaches assume that interference occurs within households but not across households, known as partial interference (Sobel, 2006). More studies need to be designed and conducted to estimate VE_I for the causal approach to find much application.

8.3 Assessing Indirect, Total, and Overall Effects

8.3.1 Cluster-randomized studies

Establishing that vaccination provides population-level effects that go beyond the direct effects in the vaccinated, such as those in Figure 8.1, can have important consequences for public health policy. The overall effect of an intervention program often is the quantity of greatest interest for policy makers, as it summarizes the public health consequences of the choice of intervention strategy if adopted in a population. Establishing that vaccination produces indirect effects in the unvaccinated can make a vaccination strategy more cost-effective. Some vaccines, such as transmission-blocking malaria vaccines, have only indirect effects. The expected magnitude of these population-level effects depends not only on the magnitude of the direct effect, but also factors related to the mixing structure of the population, transmission of the disease in the population, and the distribution of the vaccine.

If evaluating population-level effects of interventions, such as the indirect, total, or overall effects, is of interest, then a cluster-randomized study generally will be the design of choice. There are two main different kinds of cluster-randomized designs of interest for evaluating vaccines. In a parallel randomized study, clusters are identified before start of the trial, and assigned to either the vaccine of interest or a control, which could be a different vaccine, standard of care, or a placebo. All clusters stay on their assigned intervention until the end

of the study. In a stepped wedge design, the order in which clusters receive vaccination is randomized before the trial (Hussey and Hughes, 2007). The clusters where vaccination has not yet been implemented serve as control clusters until the vaccination is implemented. Thus, eventually all of the clusters receive vaccination by the end of the trial. Such trials are sometimes called phased-implementation trials.

During the 2014–2016, Ebola outbreak in West Africa, a novel form of cluster-randomized design was implemented (Ebola ca suffit, 2015; Henao-Restrepo et al., 2015). The novel ring vaccination trial of one of the Ebola vaccine candidates compared outcomes in rings of contacts and contacts of contacts around a detected case and randomized each ring to receive either immediate or delayed vaccination, such as in a stepped wedge design. The idea for the ring vaccination trial was motivated by the ring vaccination strategy around cases that led to the eradication of smallpox, but in the Ebola outbreak as part of a trial design. In the context of the public health emergency and declining incidence, the design took the trial to where the transmission was most intense. The estimated vaccine efficacy was 100 percent (95% CI 79.3, 100.0) (Henao-Restrepo et al., 2015). Although the trial reported its results as efficacy, it is more comparable to total efficacy because the vaccination within a cluster could have reduced transmission in that cluster.

Table 8.2 shows that the indirect, total, and overall effects of vaccination can be estimated using incidence rates, hazard rates, Cox proportional hazards models, or cumulative incidence. This table has been modified from earlier versions where the Cox proportional hazards model was not included for indirect, total, and overall effects with the reasoning that the baseline hazard would not be the same in clusters with different levels of coverage. However, in a cluster-randomized study, before vaccination, the expected value the baseline hazards in the clusters receiving vaccine would equal those in the clusters not receiving vaccine. Thus, its use seems potentially valid. Analysis of cluster-randomized studies needs to account for clustering using either a cluster-level random effect or a marginal model such as general estimating equations. Hayes and Moulton (2017) give details on designing and analyzing cluster-randomized trials, particularly in the context of infectious disease interventions. Because of the complex effects of vaccination, simulations can be useful for designing and planning the analysis of studies (Halloran et al., 2017).

8.3.2 Causal vaccine effects and two-stage randomization

Most of the research on cluster-randomized studies has not been done in the context of causal inference. Here we present some of the progress on evaluating the different effects of vaccination in populations using causal inference based on potential outcomes. Using potential outcomes as described in Section 8.1.2 is more complicated under interference because the potential outcomes of an individual can depend on the treatment assignment in others. If there are two treatments and a binary outcome and no interference, each individual has just two potential outcomes, and the individual causal effect is just the difference in the two. However, if there are N individuals that can potentially interfere with an individual, there are 2^N possible potential outcomes. The notation needs to allow that the potential outcomes for any individual depend on the vector of treatment assignments to other individuals with whom they potentially interfere (Rubin, 1978, 1990; Halloran and Struchiner, 1995). Individual causal estimands can be defined by the difference between the potential outcomes that depend on the treatment assignment vectors of those who interfere with that individual. However, this determination can become quite complex. The key is to define an average individual potential outcome that averages over the potential outcomes under all possible treatment assignments in the population for a particular allocation strategy (Sobel, 2006).

In the following, the population is assumed to be partitioned into groups (clusters) where interference can occur within groups but not across groups, known as partial interference

(Sobel, 2006). The goal is to obtain population-level estimates of the direct, indirect, total, and overall effects defined by comparing different treatment allocation strategies or policies. The question might be how does vaccinating 70 percent of the population with a vaccine of interest compare to vaccinating 30 percent. Hudgens and Halloran (2008) defined group- and population-level causal estimands for direct, indirect, total, and overall causal effects of treatment given two different treatment allocations. As demonstrated herein, we can define similar effects with many different levels of coverage. The potential outcomes are averaged over the individual, the group, and finally the population level. The direct effects are defined as the contrast between the population level average potential outcomes if individuals receive vaccine or control under a given strategy. The indirect effects are the contrast in population level potential outcomes if individuals receive control under one strategy and if they receive control under the other strategy. The total effects are the contrast in the population level potential outcomes if individuals receive vaccine under one strategy and if they receive control under the other strategy. The overall effect is the contrast between the marginal population level potential outcomes under one strategy and the marginal population level potential outcomes under the other strategy.

To obtain unbiased estimators of the group- and population-level causal estimands, Hudgens and Halloran (2008) proposed a two-stage randomization scheme, the first stage of randomization is at the group level, the second stage is at the individual level within groups. There are two approaches to randomization. We can randomize a fixed number of groups to each strategy, then within groups, a fixed number of individuals depending on the strategy assigned to that group. Alternatively, we can randomize each group to a strategy with a certain probability, for example with the flip of a biased coin, then similarly randomize each individual within a group to vaccine or control with a certain probability. Tchetgen Tchetgen and VanderWeele (2012) also developed estimators for the four causal estimands assuming a two-staged randomization scheme. Beyond vaccination, the approach is applicable to other situations with interference in groups of individuals where treatment can be assigned to individuals within groups. Baird et al. (2018) considered the two-stage randomized experimental design in the context of economic experiments to measure indirect/spillover effects.

To do inference about the four effects, variance estimators and confidence intervals are needed. Hudgens and Halloran (2008) made an assumption of stratified inference to develop variance estimators. Stratified inference assumes the effects of an intervention depend on what proportion of individuals in a group receive treatment, but not on exactly which ones. This reduces the complexity of the problem considerably. Tchetgen Tchetgen and VanderWeele (2012) also presented confidence limits for the four causal estimands. Liu and Hudgens (2014) derived the asymptotic distributions of estimators of the causal effects and confidence limits when either the number of individuals per group or the number of groups grows large. Some of these confidence intervals in general are narrower than those in Tchetgen Tchetgen and VanderWeele (2012).

8.3.3 Causal vaccine effects and observational studies

8.3.3.1 Inverse probability weighted estimators

Most studies are not randomized at two stages. In fact, we do not know of any vaccine studies to date that have been randomized at two stages. A study could be randomized at the individual level, at the group (cluster) level, or neither. Then the estimators described herein in general would be biased or inconsistent. For the observational setting where the treatment assignment mechanism is not known and there is no interference, propensity scores are one method to adjust the analysis to resemble results that might be obtained from a randomized

trial (Rosenbaum and Rubin, 1983). The propensity score is the probability that an individual receives a treatment assignment based on a function of the observed covariates. The propensity score can be used in different ways to adjust for measured confounders, including weighting by the inverse of the propensity score, called inverse probability weighting (IPW), or stratifying on them (Hong and Raudenbush, 2006).

Tchetgen Tchetgen and VanderWeele (2012) proposed IPW estimators of the direct, indirect, total, and overall causal effects in the presence of partial interference based on group-level propensity scores, and proved the estimators are unbiased when the group-level propensity scores are known. The estimators involve estimating mean potential outcomes by taking weighted averages of the observed responses where the weights include the inverse of group-level propensity scores. These IPW estimators can be viewed as a generalization of the usual IPW estimator of the causal effect of a treatment in the absence of interference. These could theoretically be used for studies that randomize at the group (cluster) level only, the individual level only, or neither.

Perez-Heydrich et al. (2014) used these IPW estimators to estimate the different effects of an individually randomized trial of cholera vaccination. They estimated the propensity scores in the presence of interference. They showed that the estimates were consistent and asymptotically normal, and provided variance estimators. This example provides a case study of how the direct, indirect, total, and overall effects of vaccination can be assessed and interpreted. The vaccine trial was conducted in Matlab, Bangladesh, from 1985–88. All children (2–15 yrs old) and women (>15 yrs old) were randomly assigned with equal probability to either of two killed cholera vaccines or a placebo. Although all women and children were randomized, only a subset participated in the trial. Unvaccinated individuals included eligible non-participants and placebo recipients. Vaccinated individuals included recipients of either vaccine. Of the total eligible population (N = 121,982), 49,300 women and children received two or more doses of vaccine.

The individuals lived in baris, i.e., clustered patrilineal households. Because this was an individually randomized study, neighborhoods (clusters) were defined from geo-referenced data on the baris using a clustering algorithm. In this study the geographic groups (clusters) were formed post-hoc. Thus, the level of vaccine coverage in each cluster was not randomized. The total number of groups was set to 700 for the main analysis. The analysis was based on the difference of the IPW-adjusted average outcomes in the relevant groups.

The IPW estimates of the direct, indirect, total, and overall effects of vaccination are presented in Figure 8.2. The estimates are given in units of cases of cholera per 1,000 individuals per year. The levels of vaccine coverage are denoted by α and α'. The direct effect estimates (Figure 8.2a) generally decrease with increasing α. The estimates vary from 5.3 (95% CI 2.5, 8.1) at coverage level $\alpha = 0.32$ to 0.6 (95% CI $-1.1, 2.3$) at $\alpha = 0.60$. These two inferences would lead to different conclusions about the vaccine. At the higher coverage, the vaccine does not seem to have a significant effect, illustrating the limitations of analyses that consider only direct effects when interference is present.

The indirect effect estimates are in Figure 8.2b. Here the contrast compares the incidence of cholera among unvaccinated individuals at coverage level α compared to unvaccinated individuals at coverage level α'. On the line along the diagonal where $\alpha = \alpha'$, the indirect effects estimates are zero. This result is because on the diagonal the comparison is between incidence at equal coverage. The indirect effect estimates are symmetric about the diagonal, positive on one side and negative on the other. Considering the positive estimates, the indirect effect estimates tend to increase with the difference between the two coverages. The largest estimate of the indirect effect is 5.3 (95% CI 2.6–8.0) between coverage levels 0.60 and 0.32. That is, we would expect 5.3 fewer cases of cholera per 1,000 person-years in unvaccinated individuals within neighborhoods with 60 percent coverage compared to within neighborhoods with 32 percent coverage.

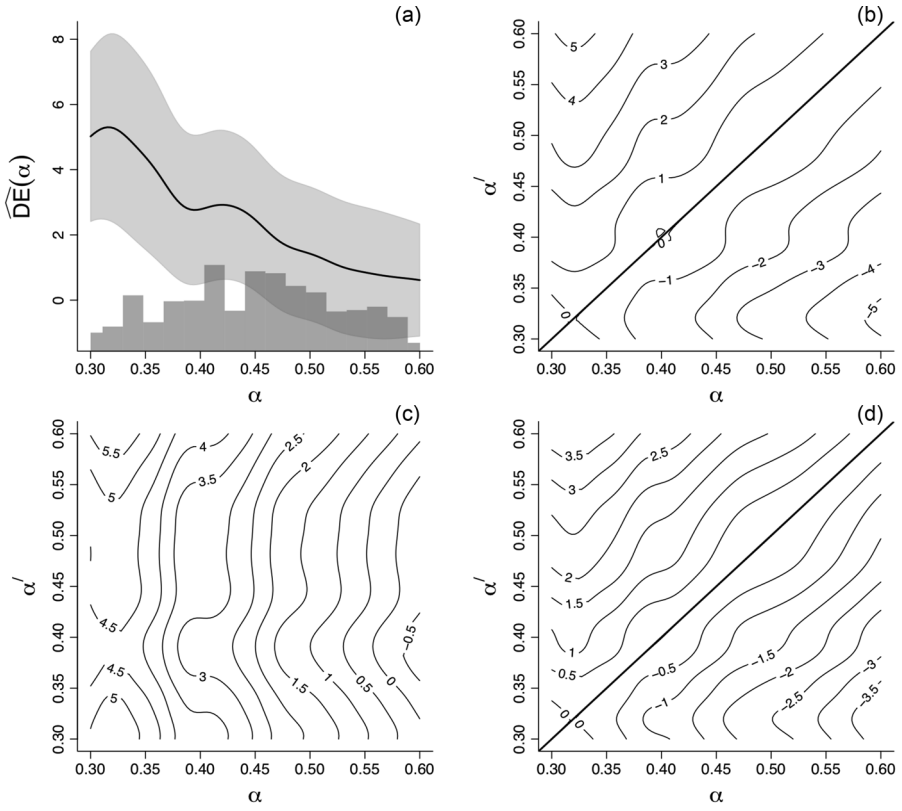

FIGURE 8.2
IPW estimates of (a) direct $\overline{DE}(\alpha)$, (b) indirect $\overline{IE}(\alpha, \alpha')$, (c) total $\overline{TE}(\alpha, \alpha')$, and (d) overall $\overline{OE}(\alpha, \alpha')$ effects based on the cholera vaccine trial data. The estimates are in units of cases per 1,000 individuals per year. In (a) the gray region represents approximate pointwise 95% confidence intervals. The histogram depicts the distribution of observed neighborhood vaccine coverage. (From Perez-Heydrich, C. et al., *Biometrics*, 70, 731–741, 2014.)

The total effect estimates in Figure 8.2c exhibit a different pattern from the indirect estimates. The total effect estimates along the line $\alpha' = \alpha$ equal the direct effects estimates in Figure 8.2a. The total effect contours are roughly vertical, suggesting the estimated risk of cholera when vaccinated tends to be the same regardless of coverage. Of particular interest is the total effect estimate of being vaccinated in a neighborhood with 60 percent coverage compared to being unvaccinated in a neighborhood with coverage 32 percent is 5.9 (95% CI 3.0, 8.8). This result is an order of magnitude greater than the estimated direct effect of 0.6 at 60 percent coverage (Figure 8.2a). Thus, taking the indirect effects into account gives a much different conclusion about the vaccine.

The estimates of overall effects (Figure 8.2d) exhibit a similar pattern to the indirect estimates. They are symmetric about the diagonal.

Details of estimation can be found in Perez-Heydrich et al. (2014). An R package `inferference` is available at CRAN that performs these analyses (Saul and Hudgens, 2017). Ali et al. (2005, 2013) found evidence of indirect effects in this cholera vaccine trial using a different, non-causal approach. To be able to estimate the population level effects of

vaccination in such studies, there must be sufficient variability in the levels of vaccine coverage. If all clusters have the same coverage, the estimates of the indirect and overall effects will all be 0, as in Figure 8.2b and d where $\alpha' = \alpha$.

Lundin and Karlsson (2014) proposed IPW estimators of direct, indirect, and total causal effects under interference where treatment assignment is randomized at the first stage but not at the second stage. They consider the situation where, in some groups, all individuals remain untreated (Figure 8.1).

IPW estimators often have relatively large variance. Liu et al. (2016) considered other forms of IPW that tend to be less variable, that is, more robust. They proposed unbiased generalized IPW estimators and two Hajek-type (stabilized) IPW estimators. One of the stabilized estimators has substantially smaller variance than the unstabilized IPW estimator proposed by Tchetgen Tchetgen and VanderWeele (2012).

8.3.3.2 One-stage cluster-randomized trial with control vaccine

To estimate indirect and total effects in cluster-randomized studies, where randomization occurs at only the cluster level, some studies use a control vaccine in the clusters not randomized to the vaccine of interest (Moulton et al., 2001). Such studies may have a target subgroup that is eligible to receive the vaccine, such as children between age 6 months and 15 years, or everyone in the population may be the target group. Not everyone in the target group randomized in a cluster to get either vaccine or control will show up, but often the cases are ascertained in the surveillance system regardless of whether they were vaccinated. Thus, estimates of the incidence in the unvaccinated individuals can still be obtained, if vaccine status is known. To estimate total effects, the average outcomes in the individuals in the vaccine clusters who show up to get the vaccine of interest are compared to those in the control clusters who show up to get the control vaccine. To estimate indirect effects, the average outcomes of those who did not get either vaccine are compared. Although these estimates will not necessarily produce the same results as the two-stage randomized design, they produce estimates of total and indirect effects based on populations assumed to be comparable by randomization. Overall effects also can be estimated.

As an example, in a cluster-randomized study of typhoid vaccine in India, Sur et al. (2009) randomized neighborhoods to receive either typhoid vaccine or control (hepatitis A) vaccine. Vaccine effects were estimated by 1 minus the relative incidence of typhoid disease in comparison groups. Total effectiveness in those who had received typhoid vaccine compared to those who had received control (hepatitis A) vaccine was estimated to be 61 percent (95% CI 41, 75). Indirect effectiveness in the unvaccinated individuals in the typhoid vaccine clusters compared to the unvaccinated individuals in the control clusters was 44 percent (95% CI 2, 69). The overall effectiveness in the typhoid clusters compared to the control clusters was 57 percent (95% CI 37, 71). However, incidence of typhoid was higher in the vaccinated individuals than in the unvaccinated individuals in the typhoid vaccine clusters (0.26 versus 0.19 cases per 100,000 person-days), and similarly in the hepatitis A vaccine clusters (0.73 versus 0.35). So if we estimated the direct effect of the typhoid vaccine in this study without any further adjustment, the estimates would be negative. The individuals who showed up to get vaccinated are clearly very different from those who did not show up to get vaccinated, either in their underlying health, health-care seeking behavior, the way they are handled in the health system for diagnosis, or other reason.

More research needs to be done on this design. As illustrated, direct effect estimates are open to extreme bias because the individuals who do not get vaccinated are likely not comparable to those who do get vaccinated, and likely should not be estimated without further consideration. The interpretation of the total and indirect effect estimates is not clear as no clear causal estimands of interest have yet been defined.

8.3.4 Surveillance or other routinely collected data

Cluster-randomized and individually randomized studies can be expensive and/or unfeasible in most situations. Surveillance data on infectious disease occurrence is gathered routinely worldwide. Large administrative databases, such as private insurance and government medical registries, also contain a trove of information on disease outcomes as well as vaccination status of individuals. Routine vaccination also occurs worldwide. It would be useful to use these available data to draw inference about population effects of vaccination (Halloran and Hudgens, 2018). Here we consider two such approaches. Both are based on comparisons of before and after introduction of a vaccination program. The first approach uses medical insurance records to estimate direct, indirect, total, and overall effects of vaccination. The second approach uses surveillance data to estimate the overall effects of introducing a vaccination program. The first approach requires individual level data on both disease outcomes and vaccination status. The second approach requires only incidence data.

Rotavirus vaccine has been available in the USA since 2006, after which the level of rotavirus vaccination in infants increased. Panozzo et al. (2014) estimated direct, indirect, total and overall effectiveness of rotavirus vaccines for the prevention of gastroenteritis hospitalizations in privately insured children in the USA in the years 2007 through 2010. Estimates of the different vaccine effects were based on Cox proportional hazards models to estimate hazard ratios of rotavirus gastroenteritis or acute gastroenteritis hospitalizations. The estimated hazard ratios in infants entering the cohort in 2007, 2008, 2009, and 2010 were used to obtain estimates by calendar year. The direct effect for each calendar year was estimated by comparing outcomes in the vaccinated and unvaccinated infants in each year from 2007 to 2010. To estimate indirect, total, and overall effectiveness during each calendar year, the comparison was made to the unvaccinated infants followed in the pre-vaccine baseline period 2001–2005. In all regression analyses, age served as the underlying time-scale. Children were censored when they experienced their first case of illness. Key to estimating all four effects was having individual-level data on outcomes and vaccination status, and a baseline pre-vaccine comparison group. The estimated direct effectiveness ranged from 87 to 92 percent, with total effectiveness 3–8 percent higher. After 4 years, indirect effectiveness reached over 70 percent and overall effectiveness was over 90 percent.

Interrupted time series analysis is a method for evaluating the overall effectiveness of a vaccination strategy that has been implemented in a population at a defined point in time. It also can be used to look at indirect effects in groups not eligible to receive the vaccine, such as age groups not targeted for vaccination, if the surveillance system has information on such groups. Lopez et al. (2017) provide a tutorial for its use for public health interventions in general. A drawback of inference based on before-and-after comparisons is that temporal trends not related to the vaccination program could reduce incidence of the disease of interest, and bias the estimate of the effectiveness of the program. The interrupted time series analysis adjusts for temporal trends occurring before introduction of the vaccination program.

The basic approach is a segmented regression. Let Y_t be the outcome at time t, X_t be a dummy variable indicating pre-vaccination (X_t set to 0) or post-vaccination period (X_t set to 1), and T is the time since the start of the study. In a simple interrupted time series analysis, the regression model is

$$E[Y_t] = \beta_0 + \beta_1 T + \beta_2 X_t + \beta_3 T X_t, \tag{8.10}$$

where β_0 is the baseline level at $T = 0$, β_1 is the change in outcome associated with time since the begin of the study, so it represents the pre-vaccination trend, β_2 is the level change after the begin of the vaccination program, and β_3 is the slope change following the intervention.

Further complications such as seasonality and other time-varying confounders can be taken into account in more complex models. See Lopez et al. (2017) for further details.

As an example, Ngabo et al. (2016) assessed the effect in Rwanda of introducing rotavirus vaccination in May 2012 into its routine national immunization schedule. Vaccine was administered at 6, 10, and 14 weeks of age. By 2013, coverage reached 99 percent in children under 1 year of age. Compared to the baseline, incidence of hospital admissions specific to rotavirus captured by active surveillance fell by 61-70 percent in the years 2012–2014. The decrease was greater in those age-eligible to be vaccinated, but there was a decrease in nearly all the age groups, suggesting a substantial indirect effect of vaccination for those not age-eligible to receive the vaccine.

8.4 Evaluating Correlates and Surrogates

An important goal of vaccine research is to identify a vaccine-induced immune response that predicts protection from the clinical endpoints of infection and/or disease. If such an immune response were known, it could be used to replace the clinical endpoint in further studies. Thus, if such an immune response were available, it would help avoid large and lengthy new trials and facilitate getting new products and formulations approved. In this section, we consider correlates and surrogates of protection. In a randomized study, the primary clinical endpoint of interest could be clinical disease, infection, or a post-infection outcome. In this section, we assume there is a binary treatment Z, such as vaccine and control, and a binary clinical outcome Y. We assume also that an inexpensive candidate immune surrogate S is measured shortly after randomization.

8.4.1 Terminology

In an important paper, Prentice (1989) defined a surrogate endpoint to be "a response variable for which a test of the null hypothesis of no relationship to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint." Prentice (1989) also proposed four criteria for a biomarker to be a surrogate endpoint for the clinical outcome of interest. In the context of vaccines (Kohberger et al., 2008), they can be stated as

1. The clinical endpoint is significantly related to the vaccine.
2. The surrogate is significantly related to the vaccine.
3. The surrogate is significantly related to the clinical endpoint.
4. The association of the surrogate with the clinical endpoint is the same in the vaccine and control groups after adjusting for baseline covariates.

For the fourth criterion to be valid, an extra assumption is needed that there are no unmeasured dual correlates of the surrogate and the clinical outcome.

Validation of a Prentice definition surrogate is a challenging task that requires much evidence, and the definition is hard to empirically validate directly (Gilbert et al., 2015). In the fourth criterion, the risk of clinical endpoints is compared in individuals with the observed values of the immunological marker. However, we observe only the immunological value and clinical endpoint that the person has under the actual vaccine assignment. We do not observe the value of the immune marker that the individual would have had under the other assignment. Thus, the fourth criteria can readily fail because of post-randomization

selection bias. The Prentice definition is clear and useful, even though the Prentice criteria can be misleading without modification (Gilbert et al., 2015).

Using the framework of potential outcomes in causal inference, Frangakis and Rubin (2002) proposed a definition of a principal surrogate based on the comparison of individuals with the same pair of potential values of the candidate surrogate under two different treatment assignments (see also Sections 8.2.6 and 8.2.7). Following that approach, Gilbert and Hudgens (2008) and Qin et al. (2007), among others, developed a framework delineating different levels of confidence in immunological markers. They particularly distinguish correlates of risk and surrogates of protection. A correlate of risk, the lowest level of confidence, is an immunological measurement that is shown to have an association with the clinical endpoint using some statistical approach. The vaccination status does not necessarily need to be taken into account.

A surrogate of protection is a correlate of risk that also predicts the level of protective vaccine efficacy. Two levels are differentiated: the specific and the general surrogates of protection. The specific surrogates can be applied in the same setting as an efficacy trial used to learn about the surrogate. The general surrogates might be applicable in other settings. Specific surrogates can be further classified as statistical surrogates of protection and principal surrogates of protection. The statistical surrogates of protection are defined in terms of the statistical and observable associations and satisfy the Prentice criteria described above. A principal surrogate is defined by strong modification of vaccine efficacy over immune marker subgroups and by the absence of vaccine efficacy in the subgroup with no vaccine effect on the immune marker.

The vaccinologist Stanley Plotkin had developed a different terminology for correlates of vaccine protection (Plotkin 2010). Plotkin and Glibert (2012) collaborated to propose a new terminology that would supplant the old. In the mapping, a specific correlate of vaccine protection would be equivalent to the specific principal surrogate of protection presented here. A bridging correlate of protection would be equivalent to the general surrogate of protection. Mechanistic and non-mechanistic correlates of protection are also differentiated. However, in the statistical literature, the terminology in Qin et al. (2007) still dominates.

8.4.2 Principal surrogates

The principal surrogates of protection are defined by fixed values of the immune response if assigned vaccine and the immune response if assigned control. The pair of potential immune responses under vaccine and control is assumed fixed before randomization to either vaccine or control, thus the pair is not subject to potential post-randomization selection bias. To begin, consider the simplest case that the potential immune response in the control would be $S(0) = 0$ or some fixed constant c. Let $S(1)$ be the potential response that an unvaccinated subject would have if vaccinated. Let Y denote having the clinical outcome ($Y = 1$) or not ($Y = 0$), and Z be assignment to vaccine ($Z = 1$) or control ($Z = 0$). Assume that the trial is randomized and that there is no interference between the units. For a specific principal surrogate of protection, one needs to estimate

$$\mathrm{VE}(s_1) = 1 - \frac{\Pr[Y = 1 | Z = 1, S(1) = s_1]}{\Pr[Y = 1 | Z = 0, S(1) = s_1]}. \tag{8.11}$$

The definition in expression (8.11) implies that the vaccine efficacy at the immune response level s_1 is one minus the relative reduction in the risk for groups of vaccinees with immune response s_1 compared with their risk if they had not been vaccinated, but if they had been vaccinated, they would have had immune response s_1. A plot of $\mathrm{VE}(s_1)$ by values of s_1 is called the causal effect predictiveness curve (Gilbert and Hudgens, 2008), or simply the vaccine efficacy curve (Huang et al., 2013). A useful principal surrogate is one where $\mathrm{VE}(s_1)$

varies greatly with values of s_1. For vaccine development, it means that increasing the level of that immune marker will increase vaccine efficacy. A good principal surrogate endpoint is relatively easy to measure and, as a strong vaccine efficacy moderator, can be used to reliably predict vaccine effects on clinical endpoints of interest. Gilbert et al. (2015) discuss the relationship of the principal stratification criteria to the Prentice definition. A good principal surrogate is not necessarily a perfect surrogate that would provide exactly the same information as the clinical endpoint, but it could still be valuable.

The problem is that in people in the control group for whom $Z = 0$, the value of s_1, the surrogate value under vaccination, is not observed. That is, s_1 is a missing potential outcome under $Z = 0$. To assess whether an immunological measurement is a specific principal surrogate of protection, knowledge about $S(1)$ is needed. That is, we need to be able to predict the immune response that an unvaccinated participant would have had if vaccinated. Follmann (2006) proposed two augmented vaccine trial designs to address the missing potential outcomes. The first approach, called "baseline immunogenicity predictor" (BIP), needs some variable measured pre-vaccination that predicts the immune response to vaccination. The BIP strategy develops an imputation model for the unobserved baseline immune measure based on the observed relationship between baseline covariates and biomarker values. Follmann (2006) suggested to vaccinate with another vaccine. But that can be hard to implement, and the immune response to two different antigens, the vaccine of interest and the irrelevant vaccine, could be quite different. One option for a good BIP is measurement of the same marker used as the candidate surrogate at baseline.

In the second approach, called "closeout placebo vaccination" (CPV), all or a subset of uninfected participants in the control group are vaccinated with the study vaccine at the end of the trial and their immune responses are recorded. An advantage in using only a subset is that by not entirely depleting the placebo group, the ability to study the durability of vaccine efficacy is retained. Then we assume that the immune response they have at the end of the trial is the response they would have had if vaccinated at the beginning of the trial. By comparing the distribution of immune responses with the full distribution of immune responses in the vaccinated group, because of randomization, we can infer what the distribution of immune responses in the infected participants in the control group would have been. Both of these approaches depend on strong assumptions when used to estimate the VE(s_1) curve.

Huang et al. (2013) proposed an efficient pseudo-score type estimator suitable for the augmented design to assess surrogate endpoint candidates in a two-phase sampling study in which the biomarker is measured in a random subcohort of study participants. Gabriel and Gilbert (2014) extended these methods to evaluating principal surrogate endpoints with time-to-event data and allowing for time-varying vaccine efficacy. The R package `pseval` (Sachs and Gabriel, 2016) on CRAN implements several functions for estimating the VE(s) curve for either a binary or a survival outcome.

8.4.3 General surrogates of protection

Although it is useful to understand the relation of immune responses to protection against infection and disease within a particular setting, the goal of identifying surrogates of vaccine protection is to replace large scale phase III trials using clinical outcomes with immunological measurements in new settings and for new vaccines. A key point is that "general" is with respect to a particular kind of bridge. It is not intended to necessarily be universal for all types of bridges. So a user must define the relevant variation in units in view of a given use of "general" For example, it could be all for one vaccine regimen, but over different pathogen serotypes, or all for one vaccine regimen and one serotype, but over different age groups. It is actually quite difficult without numerous, likely untestable assumptions.

To show that an immunological marker is a general surrogate of protection requires that it predict vaccine effects on risk across different populations, for different strains, and different vaccine products. One possible approach is to use meta-analysis (Gabriel et al., 2016) or other (Gabriel et al., 2017) approaches to combine information from several studies.

Acknowledgments

This work was partially funded by U54 GM111274, R37 AI032042, and R01 AI 085073. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH. The author is grateful to Peter Gilbert and Erin Gabriel for helpful comments.

References

Aalen O. Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann Stat*, 6:534–545, 1978.

Ali M, D Sur, YA You et al., Herd protection by a bivalent killed whole-cell oral cholera vaccine in the slums of Kolkata, India. *Clin Inf Dis*, 56:1123–1131, 2013.

Ali M, M Emch, M von Seidlein, M Yunus, DA Sack, M Rao, J Holmgren, and JD Clemens. Herd immunity conferred by killed oral cholera vaccines in Bangladesh: A reanalysis. *Lancet*, 366:44–49, 2005.

Aronow PM, NE Basta, and ME Halloran. The regression discontinuity design under interference: A local randomization approach (invited commentary). *Observational Studies*, 2:129–133, 2016.

Auranen K, E Arjas, T Leino, and AK Takala. Transmission of pneumococcal carriage in families: A latent Markov process model for binary longitudinal data. *J Am Stat Assoc*, 95:1044–1053, 2000.

Auranen K, H Rinta-Kokko, and ME Halloran. Estimating strain-specific and overall efficacy of polyvalent vaccines against pathogens with recurrent dynamics from a cross-sectional study. *Biometrics*, 69:235–44, 2013.

Baird S, JA Bohren, C McIntosh, and B Özler. Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, 2018. posted Online January 12, 2018, doi:10.1162/REST_a_00716.

Bar-Zeev N, L Kapanda, JE Tate, KC Jere, M Iturriza-Gomara, O Nakagomi et al. Effectiveness of a monovalent rotavirus vaccine in infants in Malawi after programmatic roll-out: An observational and case-control study. *Lancet Infect Dis*, 15:422–428, 2015.

Benkeser D, M Carone, and PB Gilbert. Improved estimation of the cumulative incidence of rare outcomes. *Stat Medicine*, 37:280–293, 2018. doi:10.1002/sim.7337.

Bor J, E Moscoe, P Mutevedzi, M-L Newell, and T Bärnighausen. Regression discontinuity designs in epidemiology: Causal inference without randomized trials. *Epidemiology*, 25(5):729–737, 2014.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

Cox DR. *Planning of Experiments*. John Wiley & Sons, New York, 1958.

Durham LK, IM Longini, ME Halloran, JD Clemens, A Nizam, and M Rao. Estimation of vaccine efficacy in the presence of waning: Application to cholera vaccines. *Am J Epidemiol*, 147:948–959, 1998.

Ebola ça suffit ring vaccination trial consortium. The ring vaccination trial: A novel cluster randomised controlled trial design to evaluate vaccine efficacy and effectiveness during outbreaks, with special reference to Ebola. *BMJ*, 351:h3740, 2015.

Follmann D. Augmented designs to assess immune response in vaccine trials. *Biometrics*, 62:1161–1169, 2006.

Fong Y and K Sebestyen. *kyotil: Utility Functions for Statistical Analysis Report Generation and Monte Carlo Studies*, 2018. https://CRAN.R-project.org/package=kyotil.

Frangakis CE and DB Rubin. Principal stratification in causal inference. *Biometrics*, 58:21–29, 2002.

Gabriel EE and PB Gilbert. Evaluating principal surrogate endpoints with time-to-event data accounting for time-varying treatment efficacy. *Biostatistics*, 15:251–265, 2014.

Gabriel EE, MG Sachs, and ME Halloran. Evaluation and comparison of predictive individual-level general surrogates. *Biostatistics*, 2017. doi:10.1093/biostatistics/kxx037.

Gabriel EE, MJ Daniels, and ME Halloran. Comparing biomarkers as trial-level general surrogates. *Biometrics*, 72:1046–1054, 2016. doi:10.1111/biom.12513.

Gilbert PB and MG Hudgens. Evaluating candidate principal surrogate endpoints. *Biometrics*, 64(4):1146–1154, 2008.

Gilbert PB, EE Gabriel, Y Huang, and ISF Chan. Surrogate endpoint evaluation: Principal stratification criteria and the Prentice definition. *J Causal Inf*, 3:157–176, 2015. doi:10.1515/jci-2014-0007.

Gilbert PB, V DeGruttola, MG Hudgens, SG Self, SM Hammer, and L Corey. What constitutes efficacy for a human immunodeficiency virus vaccine that ameliorates viremia: Issues involving surrogate end points in Phase 3 trials. *J Infect Dis*, 188:179–193, 2003.

Grambsch PM and TM Therneau. Proportional hazards test and diagnostics based on weighted residuals. *Biometrika*, 81:515–526, 1994.

Greenland S, JM Robins, and J Pearl. Confounding and collapsibility in causal inference. *Stat Sci*, 14:29–46, 1999.

Greenwood M and UG Yule. The statistics of anti-typhoid and anti-cholera inoculations, and the interpretation of such statistics in general. *Proc R Soc Med*, 8(part 2):113–194, 1915.

Halloran ME and CJ Struchiner. Causal inference for infectious diseases. *Epidemiology*, 6:142–151, 1995.

Halloran ME and CJ Struchiner. Study designs for dependent happenings. *Epidemiology*, 2:331–338, 1991.

Halloran ME and Hudgens MG. Estimating population effects of vaccination using large, routinely collected data. *Stat Med*, 37:294–301, 2018. doi:10.1002/sim.7392.

Halloran ME and MG Hudgens. Causal vaccine effects for infectiousness. *International J Biostat*, 8(2):1–40, 2012a.

Halloran ME and MG Hudgens. Comparing bounds on vaccine effects on infectiousness. *Epidemiology*, 23:931–932, 2012b.

Halloran ME, CJ Struchiner, and IM Longini. Study designs for different efficacy and effectiveness aspects of vaccination. *Am J Epidemiol*, 146:789–803, 1997.

Halloran ME, IM Longini, and CJ Struchiner. *Design and Analysis of Vaccine Studies*. Springer, New York, 2010.

Halloran ME, IM Longini, and CJ Struchiner. Estimability and interpretation of vaccine efficacy using frailty mixing models. *Am J Epidemiol*, 144:83–97, 1996.

Halloran ME, K Auranen, S Baird et al. Simulations for designing and interpreting intervention trials in infectious diseases. *BMC Medicine*, 15(1):223, 2017. doi: 10.1186/s12916-017-0985-3.

Halloran ME, MJ Haber, IM Longini, and CJ Struchiner. Direct and indirect effects in vaccine field efficacy and effectiveness. *Am J Epidemiol*, 133:323–331, 1991.

Hayes RJ and LH Moulton. *Cluster Randomised Trials: A Practical Approach, second edition*. Chapman and Hall/CRC, Boca Raton, FL, 2017.

Henao-Restrepo AM, IM Longini, M Egger et al. Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: Interim results from the Guinea ring vaccination cluster-randomised trial. *Lancet*, 386:857–66, 2015.

PW Holland. Statistics and causal inference. *J Am Stat Assoc*, 81:945–960, 1986.

Hong S and G Raudenbush. Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *J Am Stat Assoc*, 101:901–910, 2006.

Huang Y, PB Gilbert, and J Wolfson. Design and estimation for evaluating principal surrogate markers in vaccine trials. *Biometrics*, 69:301–309, 2013.

Hudgens MG and ME Halloran. Causal vaccine effects on binary postinfection outcomes. *J Am Stat Assoc*, 101:51–64, 2006.

Hudgens MG and ME Halloran. Towards causal inference with interference. *J Am Stat Assoc*, 103:832–842, 2008.

Hudgens MG, A Hoering, and SG Self. On the analysis of viral load endpoints in HIV vaccine trials. *Stat Med*, 22:2281–2298, 2003.

Hussey MA and JP Hughes. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*, 28:182–191, 2007.

Jackson ML and JC Nelson. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine*, 31(17):2165–2168, 2013.

Käyhty H, K Auranen, H Nohynek, R Dagan, H Mäkelä, and the Pneumococcal Carriage Group. Nasopharyngeal carriage: A target for pneumococcal vaccination. *Expert Rev Vaccines*, 5:651–667, 2006.

Kohberger RC, D Jemiolo, and F Noriega. Prediction of pertussis vaccine efficacy using a correlates of protection model. *Vaccine*, 26:3516–3521, 2008.

Liu L and MG Hudgens. Large sample randomization inference of causal effects in the presence of interference. *J Am Stat Assoc*, 109:288–301, 2014.

Liu L, MG Hudgens, and S Becker-Dreps. On inverse probability weighted estimators in the presence of interference. *Biometrika*, 103:829–842, 2016.

Lopez BJ, S Cummins, and A Gasparrini. Interrupted time series regression for the evaluation of public health interventions: A tutorial. *Int J Epidemiol*, 46:348–355, 2017.

Lundin M and M Karlsson. Estimation of causal effects in observational studies with interference between units. *Stat Methods Appl*, 23:417–433, 2014.

Mehtälä J, R Dagan, and K Auranen. Estimation and interpretation of heterogeneous vaccine efficacy against recurrent infections. *Biometrics*, 72(3):976–985, 2016.

Moscoe E, J Bor, and T Bärnighausen. Regression discontinuity designs are underutilized in medicine, epidemiology and public health: A review of current and best practice. *J Clin Epidemiol*, 68:132–143, 2015.

Moulton LH, KL O'Brien, R Kohberger, I Chang, R Reid, R Weatherholtz, JG Hackell, GR Siber, and M Santosham. Design of a group-randomised *Streptococcus pneumoniae* vaccine trial. *Contr Clin Trials*, 22:438–452, 2001.

Préziosi MP and ME Halloran. Effects of pertussis vaccination on transmission: Vaccine efficacy for infectiousness. *Vaccine*, 21:1853–1861, 2003.

Neafsey DE, M Juraska, T Bedford et al. Genetic diversity and protective efficacy of the RTS, S/AS01 malaria vaccine. *N Engl J Med*, 373(21):2025–2037, 2015.

Ngabo F, JE Tate, M Gatera, C Rugambwa, P Donnen, P Lepage, JM Mwenda, A Binagwaho, and UD Parashar. Effect of pentavalent rotavirus vaccine introduction on hospital admissions for diarrhoea and rotavirus in children in Rwanda: A time-series analysis. *Lancet Glob Health*, 4:e129–36, 2016.

Panozzo CA, S Becker-Dreps, V Pate, DJ Weber, MJ Funk, T Stürmer, and MA Brookhart. Direct, indirect, total, and overall effectiveness of the rotavirus vaccines for the prevention of gastroenteritis hospitalizations in privately insured US children, 2007-2010. *Am J Epidemiol*, 179:895–909, 2014.

Perez-Heydrich C, MG Hudgens, ME Halloran, JD Clemens, M Ali, and ME Emch. Assessing effects of cholera vaccination in the presence of interference. *Biometrics*, 70(3):731–741, 2014.

Plotkin SA and PB Gilbert. Nomenclature for immune correlates of protection after vaccination. *Clin Inf Dis*, 54(11):1615–1617, 2012.

Plotkin SA. Correlates of protection induced by vaccination. *Clin Vaccine Immunol*, 17:1055–65, 2010.

Prentice RL. Surrogate endpoints in clinical trials: Definition and operational criteria. *Stat Med*, 8:431–440, 1989.

Qin L, PB Gilbert, L Corey, MJ McElrath, and SG Self. A framework for assessing immunological correlates of protection in vaccine trials. *J Infect Dis*, 196:1304–1312, 2007.

Rhodes PH, ME Halloran, and IM Longini. Counting process models for differentiating exposure to infection and susceptibility. *J R Stat Soc Series B*, 58:751–762, 1996.

Richardson TS and JM Robins. Single world intervention graphs (SWIGs): A unification of the counterfactual and graphical approaches to causality. Technical Report 128, Center for Statistics and the Social Sciences, University of Washington, Seattle WA, 2013.

Rinta-Kokko H, R Dagan, N Givon-Lavi, and K Auranen. Estimation of vaccine efficacy against acquisition of pneumococcal carriage. *Vaccine*, 27:3831–3837, 2009.

Robins JM. A new approach to causal inference in mortality studies with sustained exposure period—Application to control of the healthy worker survivor effect. *Math Model*, 7:1393–1512, 1986.

Rodrigues L and P Smith. Case-control approach to vaccine evaluation. *Epidemiol Rev*, 21:56–72, 1999.

Rosenbaum PR and DB Rubin. The central role of the propensity score in observational studies. *Biometrika*, 70:41–55, 1983.

Ross R. An application of the theory of probabilities to the study of *a priori* pathometry, Part 1. *Proc R Soc Series A*, 92:204–230, 1916.

Rubin DB. Bayesian inference for causal effects: The role of randomization. *Ann Stat*, 7:34–58, 1978.

Rubin DB. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Stat Sci*, 5:472–480, 1990.

Rubin DB. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by Basu. *J Am Stat Assoc*, 75:591–593, 1980.

Sachs MC and EE Gabriel. *pseval: Methods for Evaluating Principal Surrogates of Treatment Response*, 2016. R package version 1.3.0. https://cran.r-project.org/web/packages/pseval/vignettes/introduction.html

Saul B and MG Hudgens. A recipe for inferference: Start with causal inference. Add interference. Mix well with R. *J Stat Softw*, 82:1–21, 2017.

Schwartz LM, ME Halloran, A Rowhani-Rahbar, KM Neuzil, and JC Victor. Rotavirus vaccine effectiveness in low-income settings: An evaluation of the test-negative design. *Vaccine*, 35:184–190, 2017.

De Serres C, DM Skowronski, XW Wu, and CS Ambrose. The test-negative design: Validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. *Euro Surveill*, 18:1–9, 2013.

Smith PG, LC Rodrigues, and PEM Fine. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. *Int J Epidemiol*, 13(1):87–93, 1984.

Sobel M. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *J Am Stat Assoc*, 101:1398–1407, 2006.

Struchiner CJ, ME Halloran, JM Robins, and A Spielman. The behavior of common measures of association used to assess a vaccination program under complex disease transmission patterns—A computer simulation study of malaria vaccines. *Int J Epidemiol*, 19:187–196, 1990.

Struchiner CJ, ME Halloran, RC Brunet, JMC Ribeiro, and E Massad. Malaria vaccines: Lessons from field trials. *Cadernos do Saúde Pública*, 10(supplement 2):310–326, 1994.

Sullivan SG, EJ Tchetgen Tchetgen, and BJ Cowling. Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness. *Am J Epidemiol*, 184:345–353, 2016.

Sur D, R Ochiai, SK Bhattacharya et al. A cluster-randomized effectiveness trial of Vi typhoid vaccine in India. *N Engl J Med*, 361:335–344, 2009.

VanderWeele TJ, E. Tchetgen Tchetgen, and M.E. Halloran. Interference and sensitivity analysis. *Stat Sci*, 29(4):687–706, 2014.

Tchetgen EJ and TJ VanderWeele. On causal inference in the presence of interference. *Stat Methods Med Res*, 21(1):55–75, 2012.

Therneau TM. *A Package for Survival Analysis in S*, 2015. version 2.38, https://CRAN.R-project.org/package=survival.

Thistlewaite DL and DT Campbell. Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *J Educ Psychol*, 51:309–317, 1960. reprinted with commentaries in Observational Studies 2:119–128, 2016.

VanderWeele TJ and EJ Tchetgen Tchetgen. Effect partitioning under interference in two-stage randomized vaccine trials. *Stat Probabil Lett*, 81:861–869, 2011.

Westreich D and MG Hudgens. Invited commentary: Beware the test-negative design. *Am J Epidemiol*, 184:354–356, 2016.