Epidemiological Reviews
Vol. 21, No. 1
Copyright © 1999 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved
Printed in U.S.A.

Design and Interpretation of Vaccine Field Studies

M. Elizabeth Halloran,[1] Ira M. Longini, Jr.,[1] and Claudio J. Struchiner[2]

INTRODUCTION

Vaccine efficacy and effectiveness (VE) are generally estimated as one minus some measure of relative risk (RR) in the vaccinated group compared with the unvaccinated group:

$$VE = 1 - RR. \quad (1)$$

Due to the dependent happenings in infectious diseases (1), vaccination can produce several different kinds of effects, at both the individual and the population levels. The groups being compared could be composed of individuals or of populations or communities. Vaccination can induce a biologically protective response in a vaccinated individual or reduce the degree or duration of infectiousness for other individuals. Widespread vaccination in a population can reduce transmission and produce indirect effects, even in individuals who were not vaccinated. Vaccinated people might change their rate of making contacts with potentially infectious sources and alter the overall benefits of vaccination. In designing a study to evaluate the effects of vaccination, the question of interest guides the choice of unit of observation, comparison groups, parameter of effect, and level of information required (2).

In this presentation we review different measures of effect of vaccination and vaccination programs. We consider various study designs for estimating the different measures of effect based on 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. Our focus is on field studies that may be double-blinded, placebo controlled phase III trials, or phase IV post-licensure studies that may be randomized or observational. We regard the paradigm of the randomized, double-blinded study as the point of departure for interpreting observational and nonrandomized studies (3, 4). Case-control studies produce approximate estimators of the appropriate population parameters (5, 6). Our primary concerns here are concepts of design and interpretation of the estimates. For example, what does it mean exactly when we say that a vaccine is 85 percent efficacious?

We begin by discussing some biologic aspects of vaccination that are of interest or that need to be taken into account when estimating vaccine efficacy. Following that, we discuss designs for estimating the direct protective effect of vaccination. Next we present design options for evaluating how vaccination alters the infectiousness of a person who becomes infected. We then consider that a vaccination may result in change in behavior, altering the exposure to infection. Continuing on, we present community-based study designs for estimating the indirect, total, and overall effectiveness of widespread vaccination in populations. We conclude by touching upon the problem of evaluating safety and adverse events.

EFFECTS OF INTEREST

Table 1 (2) provides an overview of several different types of effects and the parameters used to estimate them. Historically, the primary effect of interest of vaccination has been how well it protects the vaccinated individual. Biologically, the protective immune response can reduce the probability that a person becomes infected given a specified exposure to or inoculum of an infectious agent. That is, it can reduce the transmission probability. If a vaccinated person becomes infected, the immune response might reduce the degree or duration of disease or the probability of dying from the disease. It may also alter the rate of disease progression. For many infectious agents with

Received for publication December 10, 1998, and accepted for publication March 16, 1999.

Abbreviations: BS-WC, B subunit killed whole cell; CI, cumulative incidence; HIV, human immunodeficiency virus; *I*, infectiveness; IR, incidence rate; PH, proportional hazards; RR, relative risk; *S*, susceptibility; SAR, secondary attack rate; *T*, susceptibility and infectiveness combined effects; VE, vaccine efficacy and effectiveness; WC, whole cell.

[1] Department of Biostatistics, Rollins School of Public Health, Emory University, Atlanta, GA.
[2] Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil.
Reprint requests to Dr. M. Elizabeth Halloran, Department of Biostatistics, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322.

TABLE 1. Parameters used for measuring various effects of vaccination*

Level	Parameter choice	Comparison groups and effect: Susceptibility	Infectiousness	Combined change in susceptibility and infectiousness	
	Conditional on exposure to infection:				
I	Transmission probability, p Secondary attack rate (SAR)	$VE_{S,p}$† $= 1 - \frac{p_{01}}{p_{00}}$	$VE_{I,p} = 1 - \frac{p_{10}}{p_{00}}$	$VE_{T,p} = 1 - \frac{p_{11}}{p_{00}}$	
		Study design: I direct	IIA indirect	IIB total	III overall
	Unconditional:				
II	Incidence rate (IR)	$VE_{S,IR} = 1 - \frac{IR_{A1}}{IR_{A0}}$	$VE_{IIA,IR} = 1 - \frac{IR_{A0}}{IR_{B0}}$	$VE_{IIB,IR} = 1 - \frac{IR_{A1}}{IR_{B0}}$	$VE_{III,IR} = 1 - \frac{fIR_{A1} + (1-f)IR_{A0}}{IR_{B0}}$
	Hazard, λ	$VE_{S,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{A0}}$	$VE_{IIA,\lambda} = 1 - \frac{\lambda_{A0}}{\lambda_{B0}}$	$VE_{IIB,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{B0}}$	$VE_{III,\lambda} = 1 - \frac{f\lambda_{A1} + (1-f)\lambda_{A0}}{\lambda_{B0}}$
III	Proportional hazards (PH)	$VE_{S,PH} = 1 - e^{\beta_1}$	NA†	NA	NA
IV	Cumulative incidence (CI)	$VE_{S,CI} = 1 - \frac{CI_{A1}}{CI_{A0}}$	$VE_{IIA,CI} = 1 - \frac{CI_{A0}}{CI_{B0}}$	$VE_{IIB,CI} = 1 - \frac{CI_{A1}}{CI_{B0}}$	$VE_{III,CI} = 1 - \frac{fCI_{A1} + (1-f)CI_{A0}}{CI_{B0}}$
	Attack rates (AR)				

* From Halloran et al., Am J Epidemiol 1997;146;789–803. Reproduced with permission. 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 (see figure 3). The subscripts S, I, and T denote susceptibility, infectiousness, and combined effects, respectively. The Cox proportional hazards estimator is denoted by e^{β_1}. Time has been omitted from the table for notational clarity.

† VE, vaccine efficacy/effectiveness; NA, not applicable.

short incubation periods, disease is used as the outcome of interest in vaccine trials rather than infection. Becoming infected results with some probability from contact with an infectious source, while developing disease depends on the within host interaction subsequent to successful infection. In many vaccine studies, the distinction between infection and disease as outcome is not made. Studies with either of these outcomes are sometimes used to measure vaccine efficacy for susceptibility (VE_S) (third column, table 1), though the distinction between infection and disease should always be kept in mind.

Another measure of effect evaluates the degree of protection once a person has become infected. We call this vaccine efficacy for progression (VE_P) (not in table 1). With infectious agents that have long incubation periods, such as tuberculosis or human immunodeficiency virus (HIV), evaluation of this sort of effect is particularly important. Another example of VE_P is the comparison of the degree of illness conditional on becoming sick. For example, vaccinated persons who contract chickenpox generally have much milder disease than unvaccinated persons who contract the same disease.

The main distinction between VE_S and VE_P is that studies to estimate VE_S evaluate susceptibles and the exposure to infection would need to be taken into account. Studies to estimate VE_P are conditional on the participants already being infected, so the progression within infected individuals is important. Studies to evaluate VE_S that use disease as an outcome often do not differentiate the protective effects against infection and against disease conditional upon infection. Nonlinearities of the pathway from infection to manifest disease could mean that the efficacy as measured by the observed outcome would be quite different from the biologic efficacy if it could be measured along the pathway (7). The difference should be kept in mind when designing, analyzing, and interpreting the study.

A vaccinated person who becomes infected may also be less infectious to other susceptibles or be infectious for a shorter period of time. The vaccine efficacy for infectiousness (VE_I) (table 1, top row, columns 4 and 5), is of interest because a vaccine that reduces infectiousness could have important public health consequences. Vaccination could also change the distribution of carriers in a population.

Widespread vaccination can have indirect effects for unvaccinated people as well as for vaccinated people. The indirect effects are due to the change in collective level of immunity in the population due to vaccination. The collective level of immunity in a population against a particular parasite is called herd immunity. It is important to differentiate between the indirect effects in the unvaccinated and vaccinated groups. These are called the indirect and total effectiveness, respectively (table 1, middle, columns 4 and 5). The indirect effects in unvaccinated people might not be the same as those in the vaccinated people. The overall effectiveness of a vaccination strategy or allocation within a particular population is the weighted average of the outcomes in the vaccinated and the unvaccinated people (table 1, middle, column 6).

To evaluate the direct protective effects of vaccination, VE_S and VE_P, usually the individual is the unit of observation. To evaluate VE_I, generally small transmission units, such as households or partnerships in which contacts can be defined, are needed. This type of study in small transmission units can also be used to evaluate VE_S (table 1, top row). To evaluate the population level effects, the unit of observation becomes the population, so that several populations need to be included in the study (table 1, bottom right portion; see below in the discussion of figure 3).

Vaccination can induce heterogeneous response, and protection might wane or be boosted by exposure to wild-type infection. Exposure to infection might be heterogeneous within comparison groups. Vaccine antigen mixes might be effective against some strains, but not others. Belief in the protective effects of vaccination could result in people increasing their exposure to infection. These characteristics affect the interpretation of the vaccine efficacy estimates. In the following sections, we review some of the issues related to evaluating each of the various effects.

VACCINE EFFECT ON SUSCEPTIBILITY, VE_s

We first consider study designs for estimating the protective effects of vaccination, VE_S. In table 1, these are represented in the third column. Important for choosing methods to estimate VE_S and for interpreting VE_S estimates is to consider some model for the protection conferred by the vaccine (8). To begin, suppose that vaccination provides a multiplicative reduction in the probability of being infected given a specified exposure to infection or a specified inoculum. That is, the transmission probability in a vaccinated person is a fraction, denoted by θ, of that in an unvaccinated person. We then define $VE_S = 1 - \theta$, where VE_S represents the proportionate reduction in the transmission probability given a specified exposure to infection.

Under the assumption of equal exposure to the infectious agent in the vaccinated and unvaccinated groups (9), the estimates of VE_S are obtained from the relative risk of infection or disease in the vaccinated individuals compared with the unvaccinated individuals:

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

where R denotes one of the measures of risk. Given our assumption about the multiplicative effect of the vaccine on the transmission probability, if we can estimate the relative transmission probability to the vaccinated compared with the unvaccinated individuals, then we can estimate $VE_S = 1 - \theta$. However, this requires information on exposure to infection, which is often difficult or impossible to obtain. The incidence rate, hazard rate, and cumulative incidence (or incidence proportion) are measures of disease frequency that generally do not require knowledge of contacts with infectives to estimate. They also can be used to estimate VE_S, but then the interpretation of the estimated efficacy is different.

VE_s based on the transmission probability

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. The transmission probability could also be defined conditional on a specified level of inoculum. A related concept is the *secondary attack rate*, (SAR_{ij}), defined as the proportion of susceptibles with covariate status j making contact with an infectious person of covariate status i who become infected.

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

Then $VE_{S,p}$ based on the transmission probability or secondary attack rate (table 1, top row) is estimated from

$$VE_{S,p} = 1 - \frac{p_{.1}}{p_{.0}} = 1 - \frac{SAR_{.1}}{SAR_{.0}} =$$

$$1 - \frac{\dfrac{\text{vaccinated infections}}{\text{vaccinated exposures}}}{\dfrac{\text{unvaccinated infections}}{\text{unvaccinated exposures}}}.$$

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 within a small transmission unit such as a household or sexual partnership. For a sexually transmitted disease, the contact could be defined per sex act or per partnership. For pertussis, a contact could be defined as attending school the same day as an infectious person or as living in the same household during the entire period of infectiousness of a case. The mode of transmission of a parasite determines what types of contacts are potentially infectious.

The type of contact and the infectiousness of the infective source will determine the inoculum level per contact. A vaccine that is protective against a low inoculum might not protect against a high inoculum. If it were possible to measure the infectiousness of the infectives, the covariate j in the transmission probability might include information about this, and, therefore, provide a means to stratify by inoculum level in computing the efficacy estimates. Similarly, if it were possible to measure the different types of contacts, then the transmission probability for each type of contact could be estimated, and the $VE_{S,p}$ estimates could be stratified by type of contact. If it is not possible to measure the levels of infectiousness, the inoculum level, or the different types of contacts, then the estimates will reflect the unmeasured heterogeneities.

There are two main ways to design a study to estimate the relative transmission probabilities. The first method, called the secondary attack rate (10–13), or case-contact rate method, has been used since the pertussis vaccine trials in 1930s (14) to estimate vaccine efficacy. In this case, the focus is on identifying the infected individuals and the proportion of people exposed to them who become infected. Another method of estimating the transmission probability is based on the binomial model. In this case, we observe susceptible people, count the number of contacts they make with infectives, and count the number of these susceptible people who become infected. The transmission probability is estimated using the binomial model. Secondary attack rate studies are commonly used for directly transmitted infectious agents with high transmission probabilities, such as measles, chickenpox, mumps, pertussis, and tuberculosis. Contacts are often defined within small transmission units such as households. In tuberculosis, contact tracing and testing are often used to estimate the secondary attack rates. The binomial model is commonly used in studies with low transmission probabilities, such as HIV, in which susceptibles often make more than one contact before becoming infected. The ascertainment of the susceptibles or infectives can occur prospectively or retrospectively, depending on the design of the study.

VE_S not conditional on knowledge of exposure to infection

Information on exposure to infection is often difficult or impossible to collect. More commonly, studies are designed to estimate VE_S from events per person-time of potential rather than actual exposure or simply from the proportion of people who become infected in the vaccinated compared with the unvaccinated groups. Standard parameters for estimating VE_S are incidence rates, hazard rates, or cumulative incidence. Halloran and Struchiner (5, 15) called these comparisons study design type I (table 1, third column, lower rows). Here information on actual exposure to infection is not required to estimate VE_S. The assumption is made that the vaccinated and unvaccinated groups are equally exposed to infection (9), so that any differences in the risk in the two groups is due to the biologic effects of the vaccine.

Estimation of the different VE_S parameters requires differing levels of information and make different demands on study design and data collection (16). Incidence rates or hazard rates require the time to event and the period of potential exposure of each person under study. The hazard rate in infectious diseases is often called the *force of infection*. A Cox proportional hazards model requires only the ordering of the event times. An estimate of cumulative incidence requires only final value data, that is, whether an infection occurred by the end of the study or not. Correspondingly, in table 1, $VE_{S,IR}$ based on incidence rates (IR) and $VE_{S,\lambda}$ are level II parameters, $VE_{S,PH}$ based on Cox proportional hazards (PH) is level III, and $VE_{S,CI}$ based on cumulative incidence (CI) or final value data is level IV. Each level requires less information about the transmission system, with only level I requiring actual contact information. Thus, the levels

form a hierarchy. Since $VE_{S,p}$ based on the transmission probability is defined conditional on exposure to infection, it is called a conditional parameter, while the other measures are called unconditional parameters.

Primary vaccine efficacy studies often report $VE_{S,IR}$ based on relative events per person time, or level II information,

$$VE_{S,IR} = 1 - \frac{\text{vaccinated events/person-person-time}}{\text{unvaccinated events/person-time}}.$$

The usual assumption is that the numbers of events follow a Poisson distribution. Similarly, investigators may estimate the hazard rates in the vaccinated and unvaccinated $\lambda_1(t)$ and $\lambda_0(t)$, respectively, using survival analysis methods. Then the VE_S is based on the hazard rate ratio $VE_{S,\lambda}(t) = 1 - \frac{\lambda_1(t)}{\lambda_0(t)}$. 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 effect of the vaccine is multiplicative, constant, and homogeneous, the Cox proportional hazards model can be used. In this case, it is not necessary to estimate the hazard rate in the unvaccinated group, but only the relative hazard rate. This requires only the ordering of the infection times, as mentioned above. Covariates including time-dependent covariates can easily be incorporated using standard software.

An early example of estimating $VE_{S,IR}$ is the study by Kendrick and Eldering of pertussis vaccine in the 1930s (14), which reported both the proportion of people exposed to infection who developed pertussis, as well the number of cases per person-time. The vaccinated and control groups had 1,815 and 2,397 children, respectively, who contributed 2,268 and 2,307 person-years at risk, respectively. There were 52 cases in the vaccinated and 348 cases in the control group, so

$$\widehat{VE}_{S,IR} = 1 - \frac{\dfrac{52 \text{ cases}}{2{,}268 \text{ person-years}}}{\dfrac{348 \text{ cases}}{2{,}307 \text{ person-years}}} = 0.85. \qquad (2)$$

More recently, Urdaneta et al. (17) present estimates of $VE_{S,IR}$ as the result of a randomized, placebo controlled field trial of SPf66 malaria vaccine in Costa Marques, Rondonia, Brazil. A total of 572 participants completed the three dose vaccine schedule and were followed up for 18 months. The 287 vaccinated individuals contributed a total of 12,178 person-weeks at risk, and 76 first *Plasmodium falciparum* malaria episodes were observed among them. In the placebo group, 285 individuals contributed 11,698 person-weeks at risk and 85 cases leading to an estimate of $\widehat{VE}_{S,IR} = 0.14$.

In some studies, it is possible to compute both a conditional and an unconditional estimate of vaccine efficacy from a single study. The Kendrick and Eldering study on pertussis vaccine mentioned above also had information on children who had been exposed to pertussis within their own households. In the vaccinated group, 29 of 83 exposed children developed pertussis, while 143 of 160 exposed children in the unvaccinated group developed pertussis. Thus, the estimate of $VE_{S,p}$ is

$$\widehat{VE}_{S,p} = 1 - \frac{29 \text{ cases}/83 \text{ vac exposed}}{143 \text{ cases}/160 \text{ unvac exposed}} = 0.61. \qquad (3)$$

While everyone is included in the estimate of $VE_{S,IR}$, only the children with (presumed) exposure to infection are included in the $VE_{S,p}$ estimate. The interpretations of the two estimates are also different, since one measures the protection conferred as measured by infections per person time and the other by the probability of an infection per potentially infectious contact.

Estimation of $VE_{S,CI}(T)$ based on the cumulative incidence requires only information about whether persons are infected or not by the end of the study at time T, that is, final value data:

$$VE_{S,CI}(T) =$$

$$1 - \frac{\text{vaccinated infection events/persons-at-risk}}{\text{unvaccinated infection events/persons-at-risk}} =$$

$$1 - \frac{CI_1(T)}{CI_0(T)}.$$

As an example, Greenwood and Yule (9) used the cumulative incidence in studying the efficacy of typhoid vaccination in the troops in the early part of the twentieth century. In one analysis, Greenwood and Yule assumed that the denominators were based on the vaccinated and unvaccinated groups at the beginning of the study. They had 56 cases of typhoid in 10,378 vaccinated soldiers, and 272 cases in 8,936 unvaccinated soldiers. The estimated efficacy based on these numbers is

$$\widehat{VE}_{S,CI}(T) = 1 - \frac{56 \text{ cases}/10{,}378 \text{ at-risk}}{272 \text{ cases}/8{,}936 \text{ at-risk}} = 0.82. \qquad (4)$$

A more recent example is the estimation of $VE_{S,CI}(T)$ from a double-blinded live attenuated, cold-adapted

influenza vaccine trial in children (18). In this trial, of the 1,070 children who received vaccine, 14 were infected, and of the 532 children who received placebo, 95 were infected. Infection was defined as being culture positive for influenza virus during the acute phase of an influenza-like illness. Based on these numbers, $\widehat{VE}_{S,CI}(T) = 1 - \frac{14/1,070}{95/532} = 0.93$.

Because of the dependent happening structure of events in infectious diseases, there is an intrinsic relation among the different parameters on which the VE_S estimators are based. Understanding this relation helps to see the relation of the different estimators of VE_S to one another. Let p_{ij} be the transmission probability as defined above. 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:

$$\lambda(t) = cp_{ij}P(t). \quad (5)$$

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. Even though the cumulative incidence estimate is a sort of black-box estimator, it is useful in vaccine studies to think about the underlying transmission system that would produce the observed final values.

When estimates of VE_s vary with time

In introducing the various VE_s estimators, we made the very simple assumptions that the protective effect was the same multiplicative effect in everyone, that protection did not wane nor was it boosted, and that exposure to infection was not only the same in the vaccinated and unvaccinated groups, but that everyone was exposed equally. Unmeasured heterogeneities in susceptibility, protection, and exposure to infection can produce time-varying estimates of $VE_{S,IR}(t)$ or $VE_{S,\lambda}(t)$ that are artifacts of the choice of analysis, while true waning of protection or boosting can lead to real time-varying effects.

Consider first the situation that protective efficacy actually wanes with time. One way to build time dependence into the analysis is to partition the time axis into piece-wise constant components. Then a separate constant $VE_{S,IR}(t)$ is estimated for each time interval. If there is time dependence, then the estimates will vary across the time intervals. This method of analysis was used to estimate $VE_{S,IR}(t)$ of killed whole-cell-only (WC) and B subunit killed whole-cell (BS-WC) oral cholera vaccines over 3 years (1985–1988) of a randomized, double-blinded vaccine trial in rural Matlab, Bangladesh (19). The placebo was a killed *Escherichia coli* strain. Participants included 89,596 subjects aged 2–15 years (male and female) and greater than 15 years (females only). The first 3 years of the vaccine trial were partitioned into 1-year segments. Cholera incidence was defined as the number of culture-confirmed cholera episodes per 10^6 person-days of follow-up during each of 3 years.

In the first year of the trial, the estimated incidence rate for those vaccinated with the BS-WC vaccine was $41/10^6$ person-days and for the placebo $110/10^6$ person-days. Thus, in the first year $\widehat{VE}_{S,IR}(1) = 1 - \frac{41}{110} = 0.63$ (one tailed 95 percent confidence limit 0.50). The VE estimates for the BS-WC vaccine for the second and third year of the trial were 0.57 (one tailed 95 percent confidence limit 0.42) and 0.17 (one tailed 95 percent confidence limit –0.15), respectively. Thus, we see a waning time trend in efficacy, with no significant protection by the third year. Since the partitioning boundaries are selected at 1-year intervals, it is not clear if the waning protection is continuous or precisely at what point in time significant protection is lost. With use of a Poisson regression including covariates, the problem still remains of how to partition the time axis into piece-wise constant components. The problem can be solved by the use of survival analysis methods described in the next section.

Nonparametric estimation of time-varying vaccine effects

In this section, we assume that individual or grouped time-to-event data (level II) are available. Durham et al. (20) adapted and compared two basic approaches for the nonparametric estimation of smoothed curves for $VE_{S,\lambda}(t) = 1 - RR(t) = 1 - \frac{\lambda_1(t)}{\lambda_0(t)}$. The first is a generalized additive models approach that involves using a time-varying coefficient (21) version of the proportional hazards model assuming a Poisson model (22). It is useful for diagnostics to ascertain the shape of $\beta(t)$, but it cannot provide an estimator for $VE_{S,\lambda}(t)$. The other method uses Schoenfeld residuals (23, 24). The general idea is to fit an ordinary proportional hazards model to the data, then to compute the scaled dif-

ferences between the actual and expected covariate values at each event time, called Schoenfeld residuals. The scaled residuals are added to the coefficient from the proportional hazards model. The time-varying regression coefficient $\beta(t)$ is recovered by smoothing the rescaled Schoenfeld residuals. Conceptually, we are nonparametrically estimating the instantaneous hazard rate ratio $e^{\beta(t)}$, thus $VE_{S,\lambda}(t)$. Both methods provide a hypothesis test for the null $H_0 : \beta(t) = \beta$ for all t, that is, for no time-varying effects. The method using the Schoenfeld residuals is easy to use, provides an estimate of $e^{\beta(t)}$ on the natural scale, and allows easy incorporation of time-dependent covariates, so we recommend this approach in general.

Durham et al. (25) used the method involving Schoenfeld residuals to estimate smooth plots of the $VE_{S,\lambda}(t)$ for the two oral cholera vaccines from the cholera vaccine trial described above. Figure 1 shows the plot of the $VE_{S,\lambda}(t)$ estimates and the 95 percent confidence intervals for the two vaccines. The bending downward of the curves is indicative of waning. The p values for the hypothesis test for departures from the proportional hazards assumption are 0.008 and 0.002 for the estimated model of the WC and BS-WC vaccines, respectively. The WC vaccine gives fairly constant and significant protection, with a $VE_{S,\lambda}(t)$ of about 0.50 for the first 2.5 years of the trial, but then protection appears to wane rapidly. After 3 years of the trial (May 1988), the point estimate of the $VE_{S,\lambda}(t)$ is 0.245 and the 95 percent confidence interval covers zero. Protection from WC-BS vaccine starts out higher than that from WC vaccine, i.e., 0.713 versus 0.430, but then gradually wanes at a fairly constant rate, i.e., about 2–3 percent per month. This analysis provides a more complete description of the VE_S than that based on yearly incidence ratios described above.

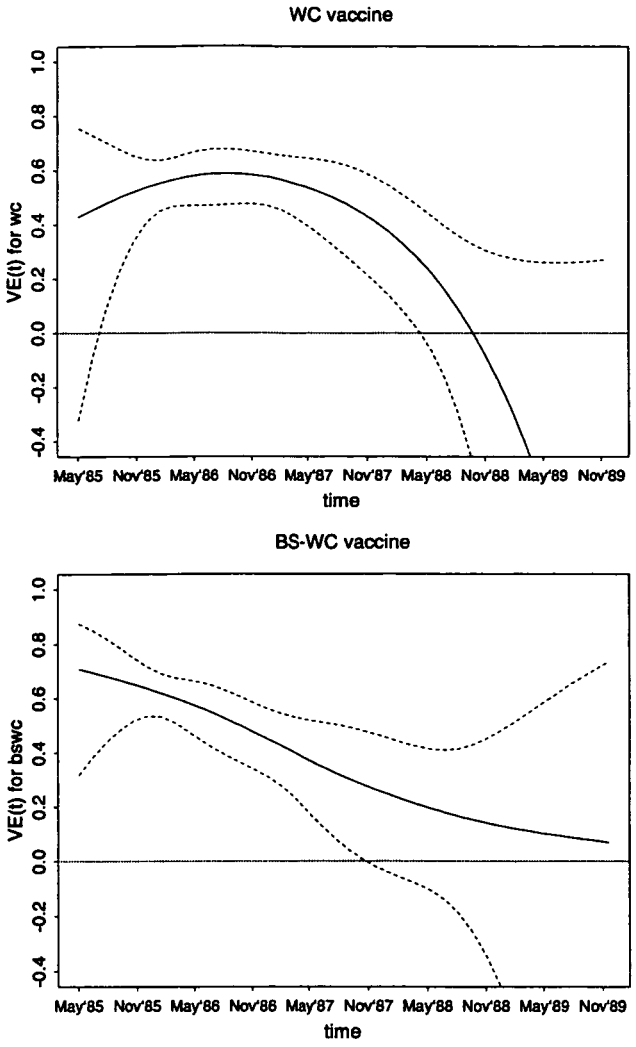

FIGURE 1. Smoothed plots of $\widehat{VE}(t)$ versus t, with 95% confidence intervals, for the whole killed cell (WC) and B subunit whole killed cell (BS-WC) vaccines, Matlab, Bangladesh, May 1, 1985, through November 31, 1989. The smoothing was carried out with regression splines with 4 degrees of freedom. These plots were constructed controlling for age as a covariate. VE, vaccine efficacy; t, time. From Durham et al., Am J Epidemiol 1998; 147:948-59. Reproduced with permission.

Unmeasured heterogeneity (frailty mixture models)

Consider now heterogeneous protection. Smith et al. (8) considered two models for vaccine protection, one which acted homogeneously and multiplicatively on the hazard and the other which protected some people completely against infection while not having any effect on the others. They showed that whether an estimator varies with time depends on the distribution of protection. We extended the discussion of mechanism of protection to the action being directly at the level of the transmission probability (26, 27). We then developed more general distributions of protection (28–31) that can include some people who are completely protected, some who have no protection, and the rest having a continuous distribution of protection. Vaccines with a multiplicative effect are called "leaky" (32), while those conferring complete protection or none at all are called "all-or-none."

With an all-or-none vaccine, the estimator $VE_{S,CI}$ does not vary with time, and provides an estimate of the proportion of the vaccinated completely protected (8). If $VE_{S,IR}$ or $VE_{S,\lambda}$ is used to estimate protection of a vaccine conferring all-or-none protection, the estimates will increase with time. This is due to the people with no protection becoming infected, so the uninfected vaccinated group remaining becomes enriched with highly protected people. Thus, the $VE_{S,\lambda}$ appears to increase with time. If a proportion of the vaccinated

group is completely protected, and another portion is only partially protected, then there are additional vaccine efficacy parameters of interest to estimate. These include α, the proportion of people completely protected, and θ, the mean reduction in susceptibility in those individuals still at least partially susceptible. The summary measure under heterogeneity is $VE_{S,SUM} = 1 - (1 - \alpha)\theta$.

Longini and Halloran (30) derived a frailty mixture model for estimating α, θ, and $VE_{S,SUM}$ when there is unmeasured heterogeneity in the vaccine effects on the host's immune response. The model is also applicable if there is heterogeneity in exposure to infection, though the interpretation of the estimates is different. The model falls into the general category of frailty models (33) employed in survival analysis, but like cure models (34), allows for a point mass at zero. It also incorporates the infection process into the baseline hazard rate. The model includes a continuous family of vaccine effect distributions ranging between the two extremes of leaky and all-or-none effects (28). The model also can be used for individual time-to-event data (35). Longini and Halloran (30) applied the methods to estimate the VE_S of measles vaccine in children from an observational study of a measles outbreak in Muyinga, Burundi. They estimated that a proportion $\alpha = 0.80$ of the vaccinated children were fully protected (95 percent confidence interval: 0.69, 0.92), but the 20 percent without full protection had a higher hazard of measles illness than did the unvaccinated children. This lead to an estimated summary $VE_{S,SUM}$ of only 0.46 (95 percent confidence interval: 0.32, 0.67). The frailty model was shown to provide an adequate fit to the measles data as indicated by a nonsignificant chi-square goodness-of-fit statistic. This validated the use of the frailty model for estimating measles vaccine efficacy in this setting.

Halloran et al. (31) explored the potential use of the above-described frailty mixture model for the estimation of VE_S over the parameter space that covers the possibilities of most vaccine studies. They showed that the parameters are identifiable under reasonable field conditions as long as there is not too much right censoring. Most importantly, they showed that the conventional VE_S estimators, i.e., proportional hazards and cumulative incidence, can be considerably biased when unmeasured heterogeneity is present. This bias is removed when the correct frailty mixture model is used.

A general strategy for estimating $VE_{S,\lambda}$

We present a general strategy for estimating $VE_{S,\lambda}(t)$ from time-to-event or incidence data. The first step is to conduct diagnostics. Then, with the help of the diagnostics, we find the best estimator of the VE_S.

Diagnostics. We begin by constructing log-minus-log plots of the Kaplan-Meier or actuarial estimates of the survival curves for the unvaccinated and vaccinated groups, such as those in figure 2a (31). These plots provide information about whether the vaccine effect is leaky, all-or-none, or a mixture. In addition, they provide some information about whether vaccine induced protection is waning. If the curves are parallel, then the effect is mostly leaky (multiplicative), and we should model the vaccine effect with a proportional hazards model. Any divergence from parallelism indicates time-varying effects and the presence of some form of heterogeneity and/or waning protection. In this case, a model other than the proportional hazards model is needed. If the curves tend to diverge, then there is all-or-none effect, and if they tend to converge, then the model still may be leaky, but with an unmeasured random effect (heterogeneity). Convergence could also indicate waning protection. Although construction of log-minus-log plots is an important first diagnostic step, they are sometimes difficult to interpret (30). If there are a sufficient number of events, a more informative plot is a smoothed hazard ratio plot of $VE_{S,\lambda}(t) = 1 - \lambda_1(t)/\lambda_0(t)$ as described above. The possible patterns associated with different vaccine effects are shown in figure 2b (31). A line with zero slope indicates a purely leaky or multiplicative effect. The researcher can construct a formal hypothesis test for zero slope (20, 24, 25).

Estimation. If there is no evidence of time-varying effects from the diagnostics, then the $VE_{S,PH} = 1 - e^{\beta}$ can be estimated by fitting a proportional hazards model. If there is evidence of time-varying effects, then the investigator should fit the full family of frailty mixture models. If these models provide an adequate fit to the data, then the estimated parameters may be, but are not necessarily, the appropriate measures of the VE_S. If there is evidence of waning or other time-varying effects not attributable to unmeasured heterogeneity, then the nonparametric estimate of $VE_S(t)$ itself will provide the best estimate. In this case, it may be possible to construct a time-dependent parametric model of the $VE_S(t)$ that would provide tighter confidence intervals than the nonparametric approach.

Interpretation of measures. Which parameter to use to estimate VE_S in a particular study depends on the type and duration of the study, the infectious agent and its transmission mode, the resources available, and the assumptions of the distribution of protection within the vaccinated group. Given the above discussion, there are clear limits on the interpretability and generalizability of estimates of VE_S. Even if time-dependent effects are detected, knowledge of the underlying biology will need to be used to interpret the effects and to

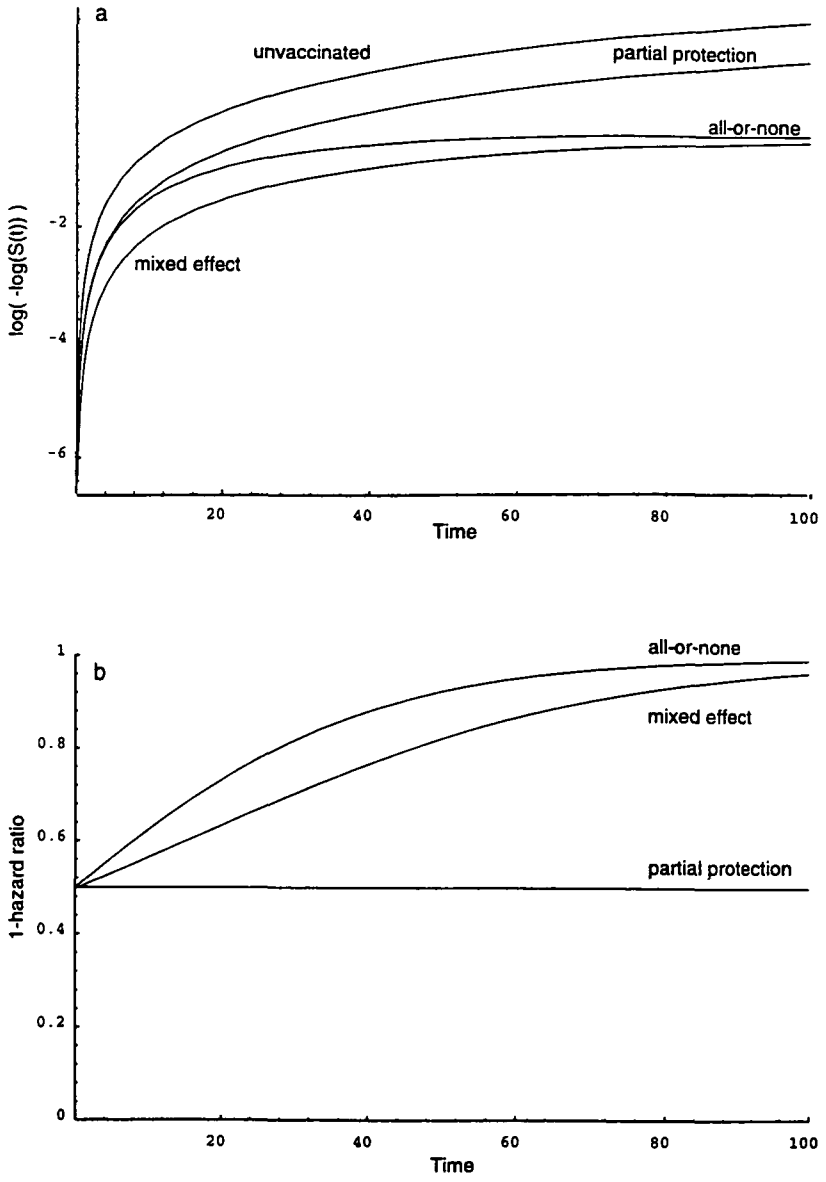

FIGURE 2. a, Diagnostic natural log ln (–ln $S_v(t)$ survival plots checking the proportional hazards assumption for a vaccine conferring homogeneous partial protection, an all-or-none vaccine, and a mixed degenerate vaccine model compared with the unvaccinated group. b, different possible 1 – hazard ratio shapes over time. Plots of 1 – hazard ratios for homogeneous partial protection ($\theta = 0.5$), the all-or-none vaccine ($\alpha_1 = 0.5$), and the mixed degenerate model ($\theta = 0.75$, $\alpha_1 = 0.33$). $VE_{SUM} = 0.5$ at time $t_0 = 0$ in these three cases. VE, vaccine efficacy; SUM, summary. From Halloran et al., Am J Epidemiol 1996;144:83–97. Reproduced with permission.

help choose between actual waning, boosting, or heterogeneities. Most vaccine trials nowadays collect data on immune response, some using a data structure of longitudinal data and repeated measures. This information can be used to help estimate and interpret vaccine effects. Also, measuring actual or potential exposure to infection in individuals will help identify heterogeneities in exposure to infection. Some trials of vaccines for vector-borne diseases have entomologic data. These help in quantifying potential exposure to infection.

That randomization does not control for confounding of estimates has been discussed by several authors in the noninfectious disease setting (36–41). Struchiner and Halloran (42) show that these results hold as well in randomized vaccine trials, but a new dimension is added when the possibly omitted covariate is exposure to infection, as is the case when going from $VE_{S,P}$ based on the transmission probability to VE_S based on the unconditional estimators. Differences in transmission intensity, previous expo-

sure to infection, and preexisting partial immunity and heterogeneities across communities result in different VE_S estimates, even when the actual biologic action of the vaccine is the same conditional on these factors. Reporting of any estimate of VE_S should include the general elements of the study design, for instance whether the estimate is based on events-per-person time, the transmission probability, or final value data. It should also include indications concerning heterogeneities or time-dependence of efficacy. However, Struchiner et al. (7) demonstrate that the complex biology and possibly nonlinear relation between the biologic mechanism of action of vaccine and the measured outcome make it important to interpret the estimates with caution. Reviews of pertussis vaccine trials in different populations using different estimators consider some of these issues (13, 43).

In addition, systems of causal inference that define causal effects in terms of potential outcomes (44, 45), that is the potential outcome if a person receives a certain treatment, say a vaccine, compared with if the person received another treatment, such as a placebo or other vaccine, generally assume that the outcomes in any individual are independent of the treatment assignment of other individuals. It is obvious that this assumption does not hold in most infectious diseases due to the dependent happenings (1). Halloran and Struchiner (46) discuss several of the open questions of causal inference in relation to vaccine studies.

Additional concerns. The definition of a case in a vaccine study is important. The sensitivity and specificity of the case definition can be crucial in determining the magnitude of the efficacy or effectiveness estimate. Often several case definitions are used in the analysis to explore how changing the case definition alters the estimates of vaccine effect.

It is possible that there are several wild-type strains of a parasite circulating in the host population. The vaccine under investigation may contain just one antigenic variant, or it may contain a cocktail of the several variants, but still not have all the wild-type variants. As molecular epidemiology and immune diagnostic measures improve, strain-specific estimates of VE_S and VE_I will be in increasing demand. Gilbert et al. (47) discuss statistical methods for inferring how VE_S may vary with viral type when several wild-type strains might be circulating in a vaccine study. The methods are presented for final value data, that is level IV, with the further additional information on each infection of what strain of wild-type parasite caused the infection. Additional considerations including adjusting for covariates, multiple events, and complex outcomes, such as malaria in which morbidity is very complex, go beyond the scope of this review.

EFFECT ON PROGRESSION, VE_P

As discussed above, the protection conferred by a vaccine if a vaccinated person does become infected is of interest. A vaccine could alter post-infection disease in several possible ways. The outcome of interest could be the probability of developing disease or of death in some specified time interval after becoming infected, the rate of progression to disease after infection, or the severity of illness after becoming infected. Evaluation of the effect of prophylactic vaccination on disease progression, VE_P, requires comparison of morbidity or mortality in vaccinated people who have become infected with that in infected unvaccinated people. Thus, infection status must be determined separately from disease status. If VE_P were evaluated by the relative rate of progression to disease, then observation of the infected individuals over time would be required. If VE_P is based on relative morbidity, then appropriate definitions of morbidity levels would be necessary. Similar to the VE_S measures discussed above, VE_P would be estimated by one minus the corresponding ratio in the vaccinated compared with the unvaccinated, including in the calculation only those people who had become infected.

VACCINE EFFECT ON INFECTIOUSNESS, VE_I

The efficacy of a vaccine in reducing infectiousness, VE_I, can be estimated epidemiologically by comparing the per-contact transmission probability from vaccinated people who become infected with the transmission probability from unvaccinated people who become infected. The relative risk comparison groups are defined according to the vaccination status of the infectious person contacting the susceptible person (46). In table 1, the VE_I estimate is shown in the second column of the top row of conditional parameters. For completeness, the third column contains the estimate of combined effect of the vaccine in reducing the transmission probability if both the infectious person and the susceptible person in the contact are vaccinated, (VE_T). If we assume that the vaccine has a multiplicative effect in reducing the transmission probability from a vaccinated infectious person to an unvaccinated person, then the efficacy for infectiousness is $VE_I = 1 - \phi$. In contrast to VE_S, which can be estimated using either conditional or unconditional parameters, the VE_I can generally be estimated using only conditional measures such as the transmission probability or secondary attack rate (46, 48–50). By making strong modeling assumptions, Longini et al. (51) suggest a method for estimating the effect of the vaccine on infectiousness using grouped time-to-event data from studies in multiple populations.

Studies for estimating VE_I can be incorporated into those for estimating $VE_{S,P}$ based on the transmission probability, if the vaccination status of the infectious person in a contact is known. The analysis can then simply stratify on the vaccination status of both the infectious and susceptible persons in the contact to get estimates of VE_S, VE_I, and VE_T. In the case of the binomial model, the likelihood can simply be constructed from the different contributions of each contact, where the parameters θ for relative susceptibility and ϕ for relative infectiousness are built directly into the likelihood (49, 52). In this case, it is simplest to assume that the effect on infectiousness and susceptibility are multiplicative and independent, so that p_{11}, = $\phi\theta p_{00}$.

Estimating VE_S and VE_I from multilevel information

In some vaccine studies, there may be information on contacts within transmission units such as households or sexual partnerships, but the individuals may also be exposed to infection outside of the transmission unit. Such an example is the pertussis study by Kendrick and Eldering (14) mentioned above. It may also be that some individuals in a study are not members of clearly defined transmission units. In these cases, it is possible to develop a statistical model to express both the within household transmission probability and the probability of being infected from the community at large (53). Rather than presenting two separate analyses for conditional and unconditional estimates, the probability model includes the probability of being infected within the transmission unit from contact with an infective and the probability of being infected outside the unit. In essence, then, the model combines elements of level I conditional parameters with elements of levels II, III, or IV unconditional parameters. People who are members of transmission units can contribute information to estimation of both VE_S and VE_I, while people who are not in transmission units and on whom no contact information is available contribute only to estimation of VE_S.

Augmented vaccine trial design. It is possible to design studies prospectively that intentionally make use of multilevel information in estimating vaccine efficacy. One such design is the *augmented* trial design (49, 52). For example, the trial may initially recruit and randomize individuals by some usual eligibility criteria. Then the trial can be augmented by including information on contacts and transmission units such as households or partnerships of the primary trial participants. This is one method to preserve the individual level analysis and randomization that some investigators think is important in randomized controlled trials. The primary analysis can still be focused on estimating VE_S, although estimation of VE_I is also possible.

Using validation sets in vaccine studies. Collecting information on contacts between infectives and susceptibles, though necessary for estimating VE_I and useful for estimating VE_S, is difficult, expensive, and inherently prone to mismeasurement. An area of current research is the possible use of validation sets for exposure to infection to improve estimation of vaccine efficacy (54–56). For example, it might be possible to get good contact information on a small subset of study participants, but only a coarse estimate on everyone else as well as the validation set. Then using methods developed for missing data (57) and two-phase sampling designs (58–62), it is possible to decrease bias and increase efficiency by combining the different levels of information.

Efficacy estimates are seriously attenuated when the diagnosis of a particular disease is not culture-confirmed but based on a nonspecific case definition. For example, if the case definition in an efficacy study of an influenza vaccine is fever and coughing, then the efficacy estimates could be much lower than the efficacy estimates based on culture-confirmed influenza. We are exploring the use of validation sets for outcomes in vaccine efficacy and effectiveness studies. In these designs, the diagnosis of the disease of interest would be confirmed by a specific test in a selected group of the study population. Statistical methods for combining validation sets with the larger main study in which the diagnosis is less specific would be used to obtain efficient, unbiased estimates of vaccine efficacy or effectiveness.

Sample size calculations. Sample size calculations should be based on the analysis method that is going to be used. For example, if the analysis is going to be based on a Cox proportional hazards model with covariates, then the sample size calculation should be based on that model, not on a model using final value data. Sample size calculations are frought with uncertainty. Even if baseline transmission intensities are measured, these may change abruptly during the trial. It is better to be generous in estimating the number of people required in the study. Precise estimation of both VE_I and VE_P require larger samples than estimates of VE_S because they both are measured on people who have already become infected, which might be a small fraction of the participants. If more complex effects such as waning, frailty models, or VE_I are to be estimated, it is good to simulate the proposed vaccine study to examine what numbers of participants are required to produce the precision desired.

CONTACT RATES AND EXPOSURE EFFICACY

Vaccinated people may alter their contact and exposure to infection patterns if they believe the vaccine is protective. *Exposure or behavior efficacy* is the relative increase or decrease in the relative risk of infection or disease due to the change in exposure to the infectious agent (54). For example, if we consider the components of the hazard rate as discussed above, changes in exposure to the infectious agent can occur in the rate of contacts, in the prevalence of infection in the contact groups, or in the transmission probability through changing the type of contact. In nonrandomized or observational studies, the vaccinated and unvaccinated groups often differ in their exposure to infection, resulting in biased estimates of VE_S. Although VE_S estimates based on the transmission probability require more information than those based on the unconditional parameters, they are less sensitive to bias from unequal exposure to infection in the two groups. The overall effect of biologic protection and change in exposure to infection might be of interest for understanding the public health consequences of vaccination. Study designs need to be explicit about differentiating factors related to susceptibility, such as vaccination status, and factors related to exposure to infection.

INDIRECT, TOTAL, AND OVERALL EFFECTIVENESS

Interest in evaluating the indirect and overall effects of vaccination strategies as part of phase III as well as post-licensure is increasing (63, 64). Struchiner et al. (5) and Halloran and Struchiner (15) define study designs for dependent happenings that allow evaluation of the indirect and overall effects of vaccination (figure 3). Since the population-level effects of a vaccine are defined within the context of a particular intervention program, or allocation of vaccination, the unit of inference is the population, and several populations or communities should be included in the study.

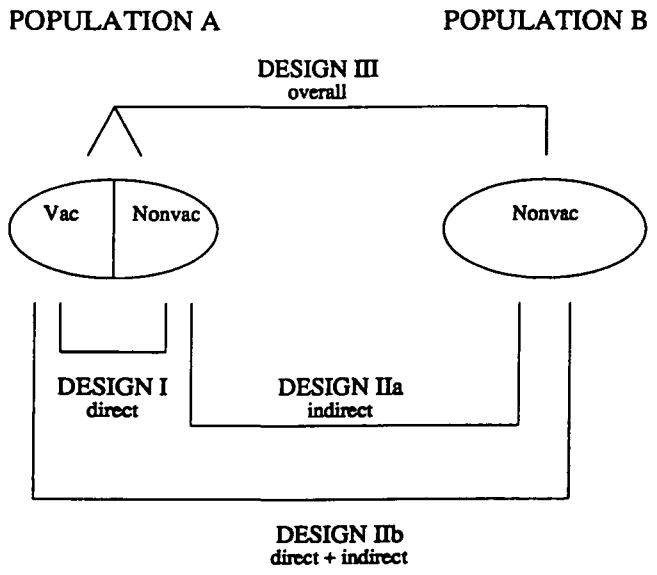

FIGURE 3. Study designs for dependent happenings. Types of effects of interventions against infectious disease, and different study designs based on comparison populations for their evaluations. From Halloran and Struchiner, Epidemiology 1991;2:331–8. Reproduced with permission.

Defining the intervention program or allocation of interest and what the comparison program or allocation of interest is can be complicated. The comparison or control populations might have no vaccination at all. The controls may be the same populations that receive the vaccination, but before the vaccination program started. In table 1 and figure 3, the different type of population level effects are considered on the simple example that no vaccination has taken place in population B, and a proportion of people are vaccinated in population A. The indirect effects of the vaccine given a specific allocation of vaccination is then the comparison of the incidence or other outcome of interest in the unvaccinated people in community A compared with the unvaccinated people in the unvaccinated community B. These comparisons are called designs type IIA. The indirect effectiveness measures are designated VE_{IIA}. The total effects of the combination of being vaccinated and the allocation is the outcome in the vaccinated people in the communities A compared with that of the unvaccinated people in the unvaccinated communities B. These comparisons are called designs type IIB, and the total effectiveness measures are designated VE_{IIB}. The overall effectiveness of the vaccine and allocation compare the average outcomes in the vaccinated communities with those of the unvaccinated communities. These comparisons are called designs type III, and the overall effectiveness measures are designated VE_{III}. Table 1 contains examples of the VE_{IIA}, VE_{IIB}, and VE_{III} based on the usual unconditional measures incidence rate, hazard rate, and cumulative incidence. Many other measures could be used, including average age of infection or the basic reproductive number, R_0.

It is important in choosing the communities or populations to assure that they are separated as much as possible in every way that is relevant for transmission. If the populations are not transmission dynamically separated, then the indirect effects of the programs will not actually differ among the groups, and the study will yield an attenuated estimate of the potential effect of intervention. Transmission patterns can differ greatly among communities. The variability could mask the effects of vaccination. Matching by transmission characteristics is an option to consider (65).

Exactly what the intervention program of interest is will depend on the vaccine and which subgroups suffer the greatest morbidity. The comparisons may be made between different levels of vaccination coverage, between allocation within different age groups or otherwise defined subgroups. For example, the intervention of interest may be vaccination of school-age children against influenza, with the primary outcome of interest the reduction in influenza incidence in unvaccinated adults. That is, the primary goal of the study might be to evaluate the pure indirect effects in unvaccinated adults. The comparison of interest would be between unvaccinated adults in populations in which children were not vaccinated and unvaccinated adults in populations with vaccinated children. If the risk measure used was cumulative incidence, then the measure would be $VE_{IIA,CI}$ in unvaccinated adults. Another outcome of interest might be the overall reduction of cumulative incidence in one influenza season in school-age children, $VE_{III,CI}$, the indirect effects in unvaccinated children of vaccinating a portion of the children, $VE_{IIA,CI}$, or the total effects in the vaccinated children, $VE_{IIB,CI}$, all compared with unvaccinated children in the unvaccinated populations.

Example

Monto et al. (66) estimated both the protective efficacy, VE_S, and the overall effect, VE_{III}, of an influenza vaccination program. They vaccinated 85 percent of the school-age children in Tecumseh, Michigan, against Hong Kong influenza just before the epidemic in 1968. The 10-week epidemic period was from November 17, 1968, to January 26, 1969. The weekly mean influenza illness rates in vaccinated and unvaccinated children were 0.072 and 0.090, respectively. This yields an approximate estimate of $VE_{S,IR} \approx 1 - \frac{0.072}{0.090} = 0.20$, which is rather low. The overall influenza illness cumulative incidence in Tecumseh for the epidemic period was 0.14, while the adjusted overall influenza cumulative incidence in unvaccinated, neighboring Adrian, Michigan, was 0.42 for the same period. Using the methods of study design III, the overall effectiveness of vaccinating 85 percent of Tecumseh's school children is estimated to be $VE_{III,CI} \approx 1 - 0.14/0.42 = 0.67$.

When designing a population-level study, it is necessary to give some thought to the likely transmission patterns and sources of exposure to infection in a population. These transmission patterns will greatly influence the magnitude of the indirect effects. For instance, if one is interested in evaluating the indirect effects in preschool children of pneumococcal vaccination, one may decide to vaccinate a large fraction of preschool children in several populations and compare it with incidence in preschool children in unvaccinated populations. However, if a large amount of the source of exposure to infection is from older children who are not included in the study, then the indirect effects of vaccination could be quite low. Vaccination of the older children as well may be necessary to have measurable indirect effects.

Because of the great variability among communities, precise definition of the actual intervention of interest will usually be difficult. When designing a study that includes several populations, the definition of the entire allocation and its implementation play an important role in the interpretability of the outcome. There will be limits in the interpretability and general applicability of the results to other settings.

Comparisons across communities would also allow study of other biologic questions. For example, vaccines might contain only particular serotypes or strains of an organism. Widespread vaccination could allow the expansion of nonvaccine serotypes that had been less important before vaccination (67, 68) or put evolutionary pressure on the existing strains. Comparison across populations or before and after introduction of vaccination would be the method to evaluate such changes.

Many issues related to the design and interpretation of studies to evaluate indirect and overall effects are active areas of research and go beyond the scope of this review. Community trials fall into the category of cluster or group randomized trials where whole social units, rather than independent individuals are randomly assigned to treatment groups (65, 69–73). The analysis and sample size calculations need to take the clustering and possible group randomization into account. Unmeasured heterogeneity may be dealt with using random effects models in the analysis phase, although such models for use in the dependent happenings setting have not yet been fully developed. Other approaches for examining population level effects include time series and population dynamic models.

Comparison with prevented fractions in noninfectious diseases

The unconditional vaccine effect parameters are analogous to the family of prevented fraction parameters discussed by Greenland and Robins (74, 75), with some essential differences. In their work, the prevented fraction in the group with a risk factor is estimated by comparing the cumulative incidence in the individuals with a risk factor to cumulative incidence in individuals without the risk factor in a study design similar to

study design I with level IV information. The group without the risk factor is supposed to represent what would have happened to the group with the risk factor had it not had the risk factor. The number of prevented cases in the group with the risk factor can be estimated from the prevented fraction based on the cumulative incidence if it is known how many people had the risk factor. Under dependent happenings, however, in study design I, say in population A, the number of cases in the unvaccinated individuals does not represent the number of cases that would have occurred in the unvaccinated individuals had the vaccination program not taken place. If vaccination is widespread enough, the cumulative incidence in the unvaccinated group will usually be lower in the presence of the vaccination program than if no one had been vaccinated. Thus, the estimated number of cases prevented generally will underestimate the actual number of cases prevented in the vaccinated group if calculated using methods for noninfectious diseases as in study design I. The comparison needs to be made between the cumulative incidence in the vaccinated group and what the cumulative incidence would have been in the unvaccinated group if no vaccination program had taken place, as in study design IIB. A similar argument applies to estimation of the prevented hazard fraction.

Multiple questions within a study

Conducting a trial to evaluate effectiveness across several different populations or communities does not preclude evaluating VE_S or VE_I of vaccination within the populations. A phase III vaccine trial can be designed to answer several questions at the same time. Randomization within a population can be used to answer efficacy questions, while comparison across populations can be used to evaluate the indirect and overall effects of vaccination. Consider a study of vaccination in several populations to measure the indirect, total, and overall effects of vaccination with different levels of coverage in each population. Within each population, a comparison can be made of the relevant vaccinated and unvaccinated portions of the population to estimate VE_{IIA}, VE_{IIB} and VE_{III}. If information is gathered within the populations on actual contacts, then the effect of the vaccine on infectiousness as well as susceptibility could be evaluated. The most important consideration in designing a vaccine study is to be clear about the effect(s) or question(s) of interest, and the level of information that can be gathered. Then the parameters of interest and the choice of comparison populations should be chosen to provide the effect measures of interest. There is a tradeoff in designing studies to measure both direct and indirect effects of vaccination between vaccinating high numbers of people so that indirect effects are high, and vaccinating too many people so that the number of events in the vaccinated populations is too low to estimate VE_S or VE_I well (51). Combining estimates of VE_S and VE_I across populations is not necessarily straightforward (42) due to lack of exchangeability (76).

SAFETY

The Food and Drug Administration in the United States is putting more emphasis on evaluating rare adverse events related to the vaccine during phase III trials. Generally, evaluation of rare adverse events has taken place after licensure when large numbers of people are vaccinated. After licensure, most studies will not be randomized, however, creating challenges for the design and interpretation of such studies. Although this brief section cannot do justice to the problem of safety evaluation, we include it as a reminder that safety evaluation belongs to all phases of vaccine studies.

An example of an adverse event is the possible association between inactivated influenza vaccine and Guillian-Barré syndrome. Since there may be a background level of such events that are not vaccine related, the rate of adverse events in the vaccinated is compared with that in the unvaccinated. If we let these adverse event rates be r_1 and r_0 in the vaccinated and unvaccinated, respectively, then the null hypothesis that the adverse events are not associated with the vaccine is $H_0 : r_1 \leq r_0$. We have the inequality in the null hypothesis since we will suspect that the vaccine is associated with the adverse events only if the adverse event rate is greater in the vaccinated. Usually the total number of adverse events is small, thus the event rate is assumed to follow a Poisson distribution. A good strategy for carrying out the above test is to condition on the total number of adverse events in both groups, and then construct the uniformly most powerful unbiased test based on the resulting binomial model (77). If the number of vaccinated and unvaccinated people at risk is reasonably small, then Fisher's exact test is used. For a large number of vaccinated and unvaccinated people, the binomial approximation is used. Several further approximations for large samples are also available (78).

In a randomized, double-blinded, phase III vaccine trial, the unvaccinated group consists of people who get placebo or some other vaccine. However, for unblinded trials, such as large community trials or post-licensure phase IV trials and surveillance, the comparison group must be carefully selected from among unvaccinated comparable people. These types of studies provide an excellent opportunity for adverse events analysis since large numbers of people are vac-

cinated, but issues concerning comparability must be addressed (79, 80).

SUMMARY

There are many different effects to consider when evaluating vaccines in the field. In this review, we have covered some of the various measures and issues related to study design and interpretation of the different measures. We emphasize that in designing and understanding vaccine studies, it is necessary to be specific about what the effect of interest is and about the assumptions underlying the interpretation of the results. Halloran et al. (81) present design, analysis, and interpretation of vaccine studies in more detail.

ACKNOWLEDGMENTS

This work was partially supported by NIH NIAID grants R01-AI32042, R01-AI40846, and the Brazilian Research Council (CNPq).

REFERENCES

1. Ross R. An application of the theory of probabilities to the study of a priori pathometry, part 1. Proc R Soc Series A 1916;92:204–30.
2. Halloran ME, Struchiner CJ, Longini IM Jr. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. Am J Epidemiol 1997;146:789–803.
3. Rosenbaum PR. Observational studies. New York, NY: Springer-Verlag, 1995.
4. Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran ME, Berry DA, eds. Statistics in epidemiology, environment and clinical trials. New York, NY: Springer-Verlag (in press).
5. Struchiner CJ, Halloran ME, Robins JM, et al. The behaviour of common measures of association used to assess a vaccination programme under complex disease transmission patterns—a computer simulation study of malaria vaccines. Int J Epidemiol 1990;19:187–96.
6. Rodrigues LC, Smith PG. Case-control approach in vaccine evaluation: efficacy and adverse effects. Epidemiol Rev 1999;21:56–72.
7. Struchiner CJ, Halloran ME, Brunet RC, et al. Malaria vaccines: lessons from field trials. Cadernos do Saúde Pública 1994;10(Suppl 2):310–26.
8. Smith PG, Rodrigues LC, Fine PEM. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. Int J Epidemiol 1984;13: 87–93.
9. Greenwood M, Yule GU. The statistics of anti-typhoid and anti-cholera inoculations, and the interpretation of such statistics in general. Proc R Soc Med 1915;8(part 2):113–94.
10. Fox JP, Hall CE, Elveback LR. Epidemiology: man and disease. New York, NY: MacMillan, 1970.
11. Orenstein WA, Bernier RH, Dondero TJ, et al. Field evaluation of vaccine efficacy. Bull World Health Organ 1985;63: 1055–68.
12. Orenstein WA, Bernier RH, Hinman AR. Assessing vaccine efficacy in the field: further observations. Epidemiol Rev 1988;10:212–41.
13. Fine PEM, Clarkson JA, Miller E. The efficacy of pertussis vaccines under conditions of household exposure: further analysis of the 1978–80 PHLS/ERL study in 21 area health authorities in England. Int J Epidemiol 1988;17:635–42.
14. Kendrick P, Eldering G. A study in active immunization against pertussis. Am J Hyg 1939;29B:133–53.
15. Halloran ME, Struchiner CJ. Study designs for dependent happenings. Epidemiology 1991;2:331–8.
16. Rhodes PH, Halloran ME, Longini IM Jr. Counting process models for differentiating exposure to infection and susceptibility. J R Stat Soc B 1996;58:751–62.
17. Urdaneta M, Prata A, Struchiner CJ, et al. Evaluation of SPf66 malaria vaccine efficacy in Brazil. Am J Trop Med Hyg 1998;58:378–85.
18. Belshe RB, Mendelman PM, Treanor J, et al. The efficacy of live attenuated, cold-adapted, trivalent, intranasal influenza-virus vaccine in children. N Engl J Med 1998;338:1405–12.
19. Clemens JD, Sack DA, Harris JR, et al. Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up. Lancet 1990;335:270–3.
20. Durham LK, Halloran ME, Longini IM Jr, et al. Smoothing methods for exploring time-dependent vaccine effects. Appl Stat 1999;48:395–407.
21. Hastie T, Tibshirani R. Varying-coefficients models. J R Stat Soc B, 1993;55:757–96.
22. Whitehead J. Fitting Cox's regression model to survival data using GLIM. Appl Stat 1980;29:268–75.
23. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239–41.
24. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994; 81:515–26.
25. Durham LK, Longini IM Jr, Halloran ME, et al. Estimation of vaccine efficacy in the presence of waning: application to cholera vaccines. Am J Epidemiol 1998;147:948–59.
26. Halloran ME, Haber M, Longini IM Jr, et al. Direct and indirect effects in vaccine field efficacy and effectiveness. Am J Epidemiol 1991;133:323–31.
27. Haber M, Longini IM Jr, Halloran ME. Measures of the effects of vaccination in a randomly mixing population. Int J Epidemiol 1991;20:300–10.
28. Halloran ME, Haber M, Longini IM Jr. Interpretation and estimation of vaccine efficacy under heterogeneity. Am J Epidemiol 1992;136:328–43.
29. Brunet RC, Struchiner CJ, Halloran ME. On the distribution of vaccine protection under heterogeneous response. Math Biosci 1993;116:111–25.
30. Longini IM Jr, Halloran ME. A frailty mixture model for estimating vaccine efficacy. Appl Stat 1996;45:165–73.
31. Halloran ME, Longini IM Jr, Struchiner CJ. Estimability and interpretation of vaccine efficacy using frailty mixing models. Am J Epidemiol 1996;144:83–97.
32. Halloran ME, Struchiner CJ, Spielman A. Modeling malaria vaccines. II: Population effects of stage-specific malaria vaccines dependent on natural boosting. Math Biosc 1989;94: 115–49.
33. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 1979;16:439–54.
34. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics 1982;38: 1041–6.
35. Haber M, Miao G. Frailty models for estimating vaccine efficacy from continuous time data. In: Bulletin of the International Statistical Institute, 51st Session, Istanbul, Turkey, August 1997:247–8.
36. Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol 1986;15:412–19.

37. Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. Am J Epidemiol 1987;125:761–8.
38. Gail MH. Adjusting for covariates that have the same distribution in exposed and unexposed cohorts. In: Moolgavkar SH, Prentice RL, eds. Modern statistical methods in chronic disease epidemiology. New York, NY: Wiley, 1986:3–18.
39. Gail MH. The effect of pooling across strata in perfectly balanced studies. Biometrics 1988;44:151–62.
40. Gail MH, Tan WY, Piantadosi S. Tests for no treatment effect in randomized clinical trials. Biometrika 1988;75:57–64.
41. Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. Biometrika 1984;71:431–44.
42. Struchiner CJ, Halloran ME. Randomization and baseline transmission in vaccine field trials. Technical report, biostatistics. Atlanta, GA: Emory University, 1996.
43. Fine PEM, Clarkson JA. Reflections on the efficacy of pertussis vaccines. Rev Infect Dis 1987;9:866–83.
44. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Stat 1978;7:34–58.
45. Holland PW. Statistics and causal inference. J Am Stat Assoc 1986;81:945–60.
46. Halloran ME, Struchiner CJ. Causal inference for infectious diseases. Epidemiology 1995;6:142–51.
47. Gilbert PB, Self SG, Ashby MA. Statistical methods for assessing differential vaccine protection against human immunodeficiency virus types. Biometrics 1998;54:799–814.
48. Koopman JS, Little RJ. Assessing HIV vaccine effects. Am J Epidemiol 1995;142:1113–20.
49. Longini IM Jr, Datta S, Halloran ME. Measuring vaccine efficacy for both susceptibility to infection and reduction in infectiousness for prophylactic HIV-1 vaccines. J Acquir Immune Defic Syndr Hum Retrovirol 1996;13:440–7.
50. Halloran ME. [Evaluating HIV vaccines] Discussion. Stat Med 1998;15:2405–12.
51. Longini IM Jr, Sagatelian K, Rida WN, et al. Optimal vaccine trial design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. Stat Med 1998;17:1121–36.
52. Datta S, Halloran ME, Longini IM Jr. Augmented HIV vaccine trial design for estimating reduction in infectiousness and protective efficacy. Stat Med 1998;17:185–200.
53. Longini IM Jr, Koopman JS. Household and community transmission parameters from final distributions of infections in households. Biometrics 1982;38:115–26.
54. Halloran ME, Longini IM Jr, Haber MJ, et al. Exposure efficacy and change in contact rates in evaluating prophylactic HIV vaccines in the field. Stat Med 1994;13:357–77.
55. Golm GT, Halloran ME, Longini IM Jr. Semi-parametric models for mismeasured exposure information in vaccine trials. Stat Med 1998;17:2335–52.
56. Golm GT, Halloran ME, Longini IM Jr. Semiparametric methods for multiple exposure mismeasurement and a bivariate outcome in HIV vaccine trials. Biometrics 1999;55:94–101.
57. Little RJA, Rubin DB. Statistical analysis with missing data. New York, NY: Wiley, 1987.
58. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73:1–11.
59. Flanders WD, Greenland S. Analytic methods for two-stage case-control studies and other stratified designs. Stat Med 1991;10:739–47.
60. Pepe MS, Fleming TR. A nonparametric method for dealing with mismeasured covariate data. J Am Stat Assoc 1991;86:108–13.
61. Reilly M, Pepe MS. A mean score method for missing and auxiliary covariate data in regression models. Biometrika 1995;82:299–314.
62. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. J Am Stat Assoc 1994;89:846–66.
63. Fine PEM. Herd immunity: history, theory, practice. Epidemiol Rev 1993;15:265–302.
64. Clemens J, Brenner R, Rao M, et al. Evaluating new vaccines for developing countries: efficacy or effectiveness? JAMA 1996;275:390–7.
65. Hayes R, Mosha F, Nicoll A, et al. A community trial of the impact of improved sexually transmitted disease treatment on HIV epidemic in rural Tanzania. 1. Design. AIDS 1995;9:919–26.
66. Monto AS, Davenport FM, Napier JA, et al. Effect of vaccination of a school-age population upon the course of an A2-Hong Kong influenza epidemic. Bull World Health Organ 1969;41:537–42.
67. Lipsitch M. Vaccination against colonizing bacteria with multiple serotypes. Proc Natl Acad Sci U S A 1997;94:6571–6.
68. Lipsitch M. Bacterial vaccines and serotype replacement: lessons fron *Haemophilus influenza* and prospects for *Streptococcus pneumoniae*. Emerg Infect Dis 1999;5:336–45.
69. Koepsell TD, Wagner EH, Cheadle AC, et al. Selected methodological issues in evaluating community-based health promotion and disease prevention programs. Annu Rev Public Health 1992;13:31–57.
70. Donner A. Some aspects of the design and analysis of cluster randomization trials. Appl Stat 1998;47:95–114.
71. Prentice RL, Sheppard L. Aggregate data studies of disease risk factors. Biometrika 1995;82:113–25.
72. Klar N, Gyorkos T, Donner A. Cluster randomization trials in tropical medicine: a case study. Trans R Soc Trop Med Hyg 1995;89:454–9.
73. Murray DM. Design and analysis of group-randomized trials. New York, NY: Oxford University Press, 1998.
74. Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fraction. Am J Epidemiol 128:1185–97, 1988.
75. Robins JM, Greenland S. Estimability and estimation of excess and etiologic fraction. Stat Med 1989;8:845–59.
76. De Finetti B. Theory of probability: a critical introductory treatment. New York, NY: Wiley, 1974.
77. Norwood PK, Sampson AR. A statistical methodology for postmarketing surveillance of adverse drug reaction reports. Stat Med 1988;7:1023–30.
78. Tsong Y. False alarm rates of statistical methods used in determining increased frequency of reports on adverse drug reaction. J Biopharm Stat 1992;2:9–30.
79. Fine PEM, Chen RT. Confounding in studies of adverse reactions to vaccines. Am J Epidemiol 1992;136:121–35.
80. Farrington CP. Relative incidence estimation from case series for vaccine safety evaluations. Biometrics 1995;51:228–35.
81. Halloran ME, Longini IM Jr, Struchiner CJ. Design, analysis, and interpretation of vaccine studies. New York, NY: Springer-Verlag (in press).