American Journal of
EPIDEMIOLOGY

Volume 146
Number 10
November 15, 1997

Sponsored by the Society for Epidemiologic Research

REVIEWS AND COMMENTARY

Study Designs for Evaluating Different Efficacy and Effectiveness Aspects of Vaccines

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

Vaccine efficacy and effectiveness (VE) are generally measured as 1 minus some measure of relative risk (RR) in the vaccinated group compared with the unvaccinated group (VE = 1 − RR). In designing a study to evaluate vaccination, the type of effect and the question of interest determine the appropriate choice of comparison population and parameter. Possible questions of interest include that of the biologic effect of vaccination on susceptibility, on infectiousness, or on progression to disease in individuals. The indirect effects, total effects, and overall public health benefits of widespread vaccination of individuals within the context of a vaccination program might also be of primary concern. The change in behavior induced by belief in the protective effects of vaccination might influence the estimates of these effects or might itself be of interest. In this paper, the authors present a framework of study designs that relates the scientific question of interest to the choice of comparison groups, the unit of observation, the level of information available for analysis, and the parameter of effect. *Am J Epidemiol* 1997;146:789–803.

clinical trials; communicable diseases; community trials; vaccine efficacy; vaccines

Vaccine efficacy and effectiveness (VE) are generally measured as 1 minus some measure of relative risk (RR) in the vaccinated group compared with the unvaccinated group:

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

Vaccination can produce several different kinds of effects, both at the individual level and at the population level, and the groups in the comparisons could be composed of individuals, populations, or communities. Vaccination can induce a biologically protective response in a vaccinated individual or reduce the degree or duration of infectiousness. 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 thereby counterbalance the biologic protective effects or alter the overall public health benefits of vaccination. In designing a study for evaluating 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.

In this paper, we present a systematic overview (table 1) of study designs for evaluating various effects of vaccination and vaccination programs based on the choice of comparison groups, the unit of observation, the choice of parameter, and the level of information required. The expression VE = 1 − RR is of the form of the prevented fraction in the exposed, and thus belongs to the family of parameters of attributable, or prevented, risk. We take as our point of departure the

Received for publication November 12, 1996, and in final form June 27, 1997.

Abbreviations: CI, cumulative incidence; IR, incidence rate; PH, proportional hazards; RR, relative risk; SAR, secondary attack rate; VE, vaccine efficacy/effectiveness.

[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, RJ 21041, Brazil.
Reprint requests to Dr. M. Elizabeth Halloran, 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)	$VE_{S,p}$† $= 1 - \frac{p_{01}}{p_{00}}$	$VE_{I,p} = 1 - \frac{p_{10}}{p_{00}}$	$VE_{T,p} = 1 - \frac{p_{11}}{p_{00}}$

Level	Parameter choice	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)				

* 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 2). 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.

review by Greenland and Robins (1, 2) of the family of parameters of attributable or prevented risk, which primarily pertained to noninfectious diseases. The two main parameters in this family that Greenland and Robins considered are the prevented hazard fraction, based on the incidence or hazard rate ratio, and the prevented fraction, based on the cumulative incidence ratio with the unit of observation being the individual. These two parameters have their analogies in the family of vaccine efficacy and effectiveness parameters. However, because of the dependent happenings (3) and the indirect effects of intervention in infectious diseases, several more measures are relevant for estimating the various prevented risk vaccine efficacy and effectiveness parameters.

In the conceptual framework we present, we expand the family of prevented risk parameters discussed by Greenland and Robins in two dimensions. In the first dimension, we add the transmission probability or secondary attack rate, a parameter that conditions on actual exposure to infection, to the incidence or hazard and cumulative incidence (4), parameters that do not condition on exposure to infection. The transmission probability is fundamental to infectious disease epidemiology. We also show that in this direction, the family of vaccine efficacy parameters form a hierarchy ordered by the amount of information required for their estimation. In the second dimension, to estimate the indirect and overall effects of widespread vaccination, we enrich the choice of comparison populations by comparing different populations or communities in the study designs for dependent happenings (5, 6). In these studies, the community or a subpopulation of the community becomes the main unit of comparison.

We also discuss the behavioral and exposure efficacies of interventions which often occur through changes in the rates and types of contacts with potentially infective sources (7). We emphasize that a study can be designed to evaluate several different types of effects at the same time. We also emphasize the importance of distinguishing risk factors for exposure to infection from risk factors for susceptibility. These concepts and study designs are applicable to risk factors and interventions other than vaccination against infectious diseases. We have found this conceptual framework useful to us in ordering our own thoughts and in discussions with other colleagues on evaluating vaccine effects. We present it here in the hope that it will provide a basis for further developments in the rational and complex design of vaccine evaluation.

BIOLOGIC EFFICACY

Figure 1 illustrates transmission between an infectious source and a susceptible person who are making

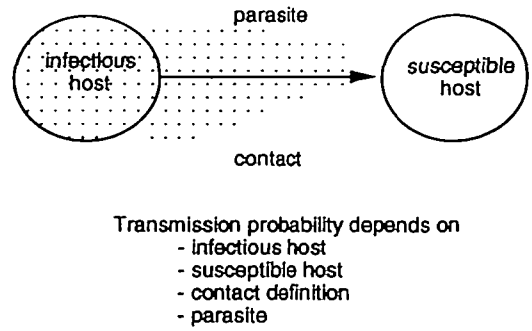

FIGURE 1. Schematic diagram of transmission of an infectious agent as an infectious host makes contact with a susceptible host.

contact with one another. The transmission probability, p_{ij}, is 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 susceptibles with covariate status j making contact with an infectious person of covariate status i who become infected. The probability of transmission depends on the characteristics of the infective source i, the infectious agent, the susceptible host j, and the type and definition of contact. For example, if 0 and 1 denote unvaccinated and vaccinated status, respectively, then p_{01} denotes the transmission probability per contact from an unvaccinated infective person to a vaccinated uninfected person. Similarly, SAR_{10} denotes the secondary attack rate from a vaccinated infective to the unvaccinated susceptibles with whom they make contact. The biologic effects of vaccination of an individual can reduce 1) the probability of infection given a specified exposure to an infectious agent, 2) the progression to, degree, or duration of disease, once the individual becomes infected, and/or 3) the degree or duration of infectiousness for other individuals. We consider estimation of the different types of effects below.

VACCINE EFFECT ON SUSCEPTIBILITY (VE_S)

The first question of interest is how vaccination protects individuals against either infection or disease. In table 1, the vaccine efficacy measures for how well the vaccine reduces susceptibility, VE_S, are given in the third column. The top row contains the conditional parameters, such as the transmission probability or secondary attack rate. The bottom rows contain the unconditional parameters incidence (hazard) rate and cumulative incidence (attack rate). Under the assumption of equal exposure to the infectious agent in the vaccinated and unvaccinated groups (8), valid

estimates of VE_S are obtained from $VE_S = 1 - [R(\text{vaccinated people})/R(\text{unvaccinated people})]$, where R denotes one of the measures of risk. We show below that the conditional and unconditional parameters form a hierarchy requiring different levels of information about the transmission system. They are also differently affected by potential sources of bias. Interpretation of the estimated VE_S as the level of biologic protection conferred by the vaccine is generally complicated, as discussed briefly below and elsewhere.

Transmission probability or secondary attack rate

An epidemiologic measure of the biologically protective effect of vaccination is the reduction in the probability of infection or disease conditional on a certain level of exposure to the infectious agent, given a certain type of contact. Thus, in the absence of controlled challenge experiments, the first choice for epidemiologically estimating the biologic protective effect of vaccination is based on the transmission probability ratio, or any similar parameter that conditions on actual contact between an infectious source and a susceptible person. The concept of a "contact" is very broad and must be defined in each particular study. In a study of sexually transmitted disease, the contact could be defined per sex act or per partnership. In a study of pertussis, a contact could be defined as attending school on 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 which types of contacts are potentially infectious.

The transmission probability is estimated in two main ways. The first method, called the secondary attack rate (9–12) or case-contact rate method, has been used to estimate vaccine efficacy since the pertussis vaccine trials were carried out in the 1930s (13). The vaccine efficacy measure is based on the $SAR_{\cdot 1}$ in the vaccinated susceptibles who were exposed to infection compared with the $SAR_{\cdot 0}$ in unvaccinated susceptibles who had a similar exposure, where the dot in the subscript can denote any vaccine status for the infective or an average across the population. Secondary attack rate studies are commonly used for directly transmitted infectious agents with high transmission probabilities, such as measles, chickenpox, mumps, pertussis, and tuberculosis. 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 binomial model is commonly used in studies with low transmission probabilities, such as studies of human immunodeficiency virus, 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.

Let $p_{\cdot 0}$ and $p_{\cdot 1}$ denote the probability of transmission to unvaccinated and vaccinated susceptibles, respectively. Analogously, let $SAR_{\cdot 0}$ and $SAR_{\cdot 1}$ denote the secondary attack rates for unvaccinated and vaccinated susceptibles, respectively. The dot denotes the infectious contacts, which are assumed to be equal for the vaccinated and unvaccinated susceptibles. $VE_{S,p}$, based on the transmission probability or secondary attack rate, is estimated from

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

$$= 1 - \left(\frac{\text{vaccinated infections}}{\text{vaccinated exposures}} \div \frac{\text{unvaccinated infections}}{\text{unvaccinated exposures}}\right).$$

Estimating vaccine efficacy from the transmission probability ratios requires information on who is infectious and when, and on whom they contact and how. Because this method generally requires the most information, Rhodes et al. (14) call the conditional measure a level I parameter (table 1, top row).

Hierarchy of VE_S parameters

Gathering information on exposure to infection and on contacts between infective and susceptible people in a vaccine study is often expensive, difficult, or even impossible. Thus, quite often vaccine efficacy studies rely on time-to-infection or time-to-disease data or simply on final value data. The analysis may be stratified according to variables believed to correlate with exposure to infection. If time-to-event data are collected, VE_S can be estimated from the incidence or hazard rate ratio; Rhodes et al. (14) called this level II information. For the Cox proportional hazards model (15), a special hazard rate ratio that assumes the same baseline hazard in both the vaccinated and unvaccinated groups is used. Then only the ordering of the events is needed; this is called level III information. If only final value data, called level IV data, are collected, VE_S is based on the cumulative incidence or attack rate ratio.

The transmission probability, incidence rate, and cumulative incidence have a fundamental relation to one another because of the dependent happening struc-

ture of infectious diseases. The incidence rate and cumulative incidence can be thought of as functions of the underlying transmission process, even if we do not measure the components. The components include the rate of contacts in the population, the probability that a contact between an infective and a susceptible will result in transmission, and the probability that any contact a susceptible makes is with an infectious source. Quite simply, the number of infection events per person-time equals the number of contacts per unit of time × the transmission probability per infectious contact × the probability that the contact is infectious. If c_0 is the baseline contact rate, $p_{.0}$ is the average transmission probability from an infectious person to an unvaccinated susceptible person, and $P_0(t)$ is the prevalence of infection in people with whom the unvaccinated people make contact at time t, the hazard $\lambda_0(t)$ or incidence rate $IR_0(t)$ in the unvaccinated population can be expressed as $\lambda_0(t) = IR_0(t) = c_0 p_{.0} P_0(t)$. The term $P_0(t)$ could also represent the probability that an environmental exposure is contaminated with the infectious agent, such as cholera bacteria in drinking water. In the vaccinated group, the incidence $IR_1(t)$ or hazard rate $\lambda_1(t)$ of infection is the product of the contact rate in the vaccinated group, c_1, the average transmission probability from an infectious person to a vaccinated susceptible, $p_{.1}$, and $P_1(t)$, the prevalence of infection in people with whom the vaccinated people make contact at time t: $\lambda_1(t) = c_1 p_{.1} P_1(t)$.

If the contact rates in the two groups are equal and the prevalences of infection and infectiousness in the groups with whom they mix are equal, perhaps through randomization, then vaccine efficacy based on the incidence $VE_{S,IR}(t)$ or hazard rate ratio $VE_{S,\lambda}(t)$ can be thought of as

$$VE_{S,\lambda}(t) = 1 - \frac{\text{vaccinated events/person-time}}{\text{unvaccinated events/person-time}} = 1 - \frac{\lambda_1(t)}{\lambda_0(t)} = 1 - \frac{c_1 p_{.1} P_1(t)}{c_0 p_{.0} P_0(t)} \cong 1 - \frac{p_{.1}}{p_{.0}}.$$

Generally, simple cancellation of the contact rates and the prevalence of infection is not possible because of heterogeneities in the vaccinated and unvaccinated groups. Therefore, estimation of vaccine efficacy from time-to-event data would not generally yield the same estimate as that based on the transmission probability. The estimated *epidemiologic* efficacy often differs from the *biologic* efficacy measure of interest. However, it is useful to think of the transmission probability and incidence rates as intrinsically related to each other. In the Cox proportional hazards model (PH), the estimate is based on the partial likelihood and requires data on the order of events, $VE_{PH} = 1 - e^{\beta_1}$, where e^{β_1} would be estimated by the hazard ratio (14).

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.

$$VE_{S,CI}(T) = 1 - \frac{\text{vaccinated infection events/persons-at-risk}}{\text{unvaccinated infection events/persons-at-risk}} = 1 - \frac{CI_1(T)}{CI_0(T)} = 1 - \frac{1 - \exp(-\int_0^T \lambda_1(t)dt)}{1 - \exp(-\int_0^T \lambda_0(t)dt)}.$$

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.

Example

It is possible to combine different levels of information within the same study by the use of appropriate probability models or by collection of different levels of data. Sometimes both a conditional and an unconditional estimate of vaccine efficacy are given from the analysis of a single study. For those study subjects who had a putative exposure to the infectious agent, the conditional estimate is given. The unconditional estimate using all eligible study subjects is also provided. An early example is the study of pertussis vaccine carried out in the 1930s (13), which reported both the proportion of people exposed to infection who developed pertussis and the number of cases per person-time. The vaccinated and control groups had 1,815 and 2,397 children, respectively, who contributed 4,575 and 2,268 person-years at risk, respectively. There were 52 cases in the vaccinated group and 348 cases in the control group, so

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

The study 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,

$$\widehat{\mathrm{VE}}_{S,p} = 1 - \frac{29 \text{ cases/83 vaccinated exposed}}{143 \text{ cases/160 unvaccinated exposed}} = 0.61. \quad (3)$$

Fine et al. (12) review several studies of pertussis vaccines and discuss some possible biologic reasons for differing values of the conditional estimates versus the unconditional estimates.

Example

Greenwood and Yule (8) used the cumulative incidence, or attack rates, in studying the efficacy of typhoid vaccination among troops in the early part of the 20th century. In one of their studies, use of the attack rates was complicated because vaccine continued to be administered to the troops until the end of the study. Thus, the denominators for the vaccinated and unvaccinated groups were changing over the course of the study. Greenwood and Yule provided two analyses, one assuming that all of the denominators were based on the vaccinated and unvaccinated groups at the beginning of the study and the other based on who was vaccinated or unvaccinated at the end of the study. Here we show the analysis using denominators based on the beginning of the study. Greenwood and Yule 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

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

Example

We compare the conditional estimators with the unconditional estimators in a simple example. Consider a 1-year, randomized, double-blinded, placebo-controlled trial with 5,000 people each in the vaccinated and unvaccinated arms of the study. Suppose that the incidence rate in the unvaccinated group is 2.0 percent, and that the vaccine reduces susceptibility to infection equally and multiplicatively in all vaccinated persons by 0.50. Ignoring indirect effects, and assuming only one contact per person (16), after 1 year the expected data would be those shown in table 2. After 1 year ($t = 1$), the estimated efficacy is

$$\mathrm{VE}_{S,\lambda}(1) = \mathrm{VE}_{S,\mathrm{CI}}(1) = 1 - \frac{50/5{,}000}{100/5{,}000} = 0.50.$$

If we were able to identify all contacts with infectives, as in the hypothetical study shown in table 3, we could estimate vaccine efficacy using the transmission probability. Under randomization and double-blinding, we assume that the number of potentially infective contacts in each group is the same. The estimated vaccine efficacy based on the transmission probability would be

$$\mathrm{VE}_{S,p} = 1 - \frac{50/1{,}000}{100/1{,}000} = 0.50.$$

Complex considerations

We chose this simple example to demonstrate the relations among the different levels of parameters. Generally things are not this simple. By conditioning on contacts with infectives, comparison of the transmission probabilities to vaccinated susceptibles versus unvaccinated susceptibles is the most likely of the vaccine efficacy measures to give meaningful information about the biologic protective effect of the vaccine. Estimates of vaccine efficacy based on the incidence rate ratio or cumulative incidence ratio can be interpreted under certain assumptions about equal exposure to infection in the comparison groups (8), how the vaccine works, and the distribution of its effects in the population (16–20). Even under the assumption of

TABLE 2. Expected data from a study collecting information on person-time at risk*

Status	No. of participants	No. of infections	Time (years)
Vaccinated	5,000	50	1
Unvaccinated	5,000	100	1

* The incidence rate in the unvaccinated group is 2.0 percent per year.

TABLE 3. Expected data from a study collecting information on number of exposures to an infectious agent

Status	No. of participants	No. of exposures to infection	No. of infections	Time (years)
Vaccinated	5,000	1,000	50	1
Unvaccinated	5,000	1,000	100	1

equal exposure to infection in the vaccinated and unvaccinated groups, $VE_{S,\lambda}$ and $VE_{S,CI}$ can give very biased estimates of the effect of vaccination on the transmission probability (17, 19–21).

Several complicating factors in the evaluation of vaccine efficacy go beyond the four levels of information and their respective parameters presented here. Some vaccines made of live, attenuated infectious agents can spread from the vaccinees to people who have not been vaccinated, a phenomenon referred to as "contagious treatment" (6). If the vaccine virus spreads to a substantial number of unvaccinated study subjects, the efficacy of the vaccine could be underestimated.

Another complicating factor is how to interpret time-varying vaccine efficacy estimates. Vaccine efficacy estimates can vary over time because of heterogeneity in susceptibility, rates of exposure to infectious agents, or the protection conferred by the vaccine (17, 19, 20). They could also vary because the efficacy of the vaccine actually wanes over time (22). The efficacy could also increase with time if the level of efficacy depends on boosting by natural infection (6). Common to all of these effects is the fact that they can be captured only if time-to-event data are available. Although heterogeneities can be taken into account with frailty models (19, 20) and waning can be estimated with parametric or nonparametric models (22), designing studies to estimate changes in efficacy due to natural boosting remains a challenge. Another open challenge is that of distinguishing among the possible causes of time-varying estimates. The overall, long-term effects of vaccination in a population depend on these characteristics (23–25), so study designs and methods of analysis for evaluating them are important areas for further development.

Practical choice of estimators

The amount of information gathered—that is, from the most (level I) to the least (level IV)—will be determined by many practical considerations that we cannot consider in detail here. The choice of estimators can be driven by the type and duration of the study. Postlicensure evaluation of a vaccine after an acute outbreak of an infectious agent often must be carried out using only information on whether people became infected or developed the disease during the period of the outbreak, since time-to-event or exposure data are not available. In this case, estimation of vaccine efficacy will be based on the cumulative incidence, or attack rates. If time-to-event data were available, relative incidence rates or relative hazards could be used. If the study occurred over a longer period of time, with the risk set also changing over time, then the relative incidence or relative hazard rates would be a more appropriate choice for the dynamic cohort that would be under observation.

CONTACT RATES AND EXPOSURE EFFICACY

Knowledge of being vaccinated could alter the contact rate of a vaccinated person or the contact pattern in a vaccinated population. With interventions other than vaccines, the primary goal of an intervention may be change in exposure to infection. "Contact rate efficacy" is the relative change in the contact rates due to an intervention. "Exposure efficacy" or "behavior efficacy" is the relative increase or decrease in the risk of infection or disease due to the change in exposure to the infectious agent, depending on the outcome measure chosen (7). 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 contact, in the prevalence of infection in the contact groups, or in the transmission probability through a change in the type of contact. In nonrandomized or observational studies, the vaccinated and unvaccinated groups often differ in their exposure to infection without changing their behavior. Unequal exposure to infection in the two comparison groups can bias estimates of the efficacy measures of interest. Although vaccine efficacy 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.

Example (continued)

Continuing the example from above, suppose that the study were randomized but not blinded and that the vaccinated individuals all doubled their rate of contact by a factor of 2, so that $c_1 = 2c_0$. The expected data from such a study are shown in table 4. After 1 year ($t = 1$), the unconditional estimated efficacy would be

$$VE_{S,\lambda}(1) = VE_{S,CI}(1) = 1 - \frac{100/5{,}000}{100/5{,}000} = 0.$$

TABLE 4. Expected data in an unblinded, randomized study in which all vaccinated individuals double their contacts by a factor of 2 after randomization

Status	No. of participants	No. of exposures to infection	No. of infections	Time (years)
Vaccinated	5,000	2,000	100	1
Unvaccinated	5,000	1,000	100	1

The efficacy estimate based on the number of infections conditional on the number of exposures to infection would be

$$VE_{S,p} = 1 - \frac{100/2{,}000}{100/1{,}000} = 0.50.$$

Using the unconditional estimates, with no knowledge of the change in exposure to infection, we would believe the protective effect of the vaccine to be zero, while, using the conditional estimates, we correctly estimate the protective efficacy to be 0.50. A similar result would occur if vaccinated people made contact with people who were more infectious or with groups in which the prevalence of infection was twice as high. Such a result would also occur if the vaccinated people changed the type of contact they made to one that had a higher baseline transmission probability.

The combined effect of the change in exposure to infection and the biologic protective effect of the vaccine is an important public health measure of interest. Indeed, an increase in the contact rate or other source of increased exposure to infection could outweigh any biologic protection conferred by the vaccine. Thus, a biologically protective vaccine could have detrimental public health effects. Whether the biologic protective efficacy or overall efficacy for an individual is of interest depends on the question underlying the study. The study design must be chosen to ensure that the question or questions of interest will be answered. In general, it is important to differentiate risk factors for exposure to infection from risk factors for biologic susceptibility. Stratification by surrogates for amount of exposure to infection could help to reduce potential bias in unconditional estimators.

It is possible to design studies to examine several questions at the same time and to estimate more than one parameter from the data. In the last example, if the number of exposures to infection were ascertained, then both $VE_{S,p}$ and $VE_{S,\lambda}$ could be estimated, giving an estimate of both the biologic protective efficacy and the effect of the combined reduction in susceptibility with increased exposure.

VACCINE EFFECT ON INFECTIOUSNESS (VE_I)

The biologic effect of the 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. In contrast to VE_S, which is estimated using either conditional or unconditional parameters, VE_I can generally be estimated only using conditional measures such as the transmission probability or the secondary attack rate (4, 26–28). To estimate the biologic effect of the vaccine in reducing infectiousness (4), the relative risk comparison groups are defined according to the vaccination status of the infectious person contacting the susceptible person. 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 of parameters contains the estimate of the combined effects of the vaccine in reducing the transmission probability if both the infectious person and the susceptible person in the contact are vaccinated (VE_T).

Example

Suppose that the goal of a study were to estimate the efficacy of a pertussis vaccine in reducing susceptibility and infectiousness. In addition to the number of potentially infectious contacts and the vaccine status of the exposed susceptibles, we also record the vaccine status of the infectious contact, possibly the primary case in the household. The data from such a hypothetical study are shown in table 5. For example, there are 44 infections in vaccinated people who had a total of 800 exposures to unvaccinated infective people. The estimate of SAR_{01} is 44/800. The effect of the vaccine in reducing susceptibility and infectiousness and the combined effects on both are estimated by

$$VE_{S,p} = 1 - \frac{SAR_{01}}{SAR_{00}} = 1 - \frac{44/800}{88/800} = 1 - \frac{SAR_{11}}{SAR_{10}} = 1 - \frac{6/200}{12/200} = 0.50;$$

$$VE_{I,p} = 1 - \frac{SAR_{10}}{SAR_{00}} = 1 - \frac{12/200}{88/800} = 1 - \frac{SAR_{11}}{SAR_{01}} = 1 - \frac{6/200}{44/800} = 0.45;$$

$$VE_{T,p} = 1 - \frac{SAR_{11}}{SAR_{00}} = 1 - \frac{6/200}{88/800} = 0.73.$$

TABLE 5. Expected data from a study using the household secondary attack rate to estimate both vaccine efficacy for susceptibility (VE_S) and vaccine efficacy for infectiousness (VE_I)

Status	No. of participants	Infectious contact	No. of exposures to infection	No. of infections	Time (years)
Vaccinated	5,000	Vaccinated	200	6	1
		Unvaccinated	800	44	
Unvaccinated	5,000	Vaccinated	200	12	1
		Unvaccinated	800	88	

This is interpreted to mean that vaccination reduces the per-contact susceptibility by 0.50 and reduces infectiousness (once a vaccinee becomes infected) by 0.45 percent, and the combined reduction in susceptibility and infectiousness reduces the transmission probability by 0.73.

Generally, contact information is required to estimate the effect of the vaccine in reducing infectiousness. By making strong modeling assumptions, Longini et al. (29) suggest a method for estimating the effect of the vaccine on infectiousness using time-to-event data from studies in multiple populations.

Effect on progression (VE_P)

Evaluation of the effect of prophylactic vaccination on progression to disease, VE_P, requires comparison of morbidity in vaccinated people who have become infected with that in infected unvaccinated people. Estimation of the effect of vaccination on progression would generally require observation of the infected individuals over time. Problems of interval censoring are important, because often the time of infection would not be observed accurately.

INDIRECT AND TOTAL EFFECTS (VE_{IIA} AND VE_{IIB})

While the parameters in the third column of table 1 provide estimates of the protective efficacy for individuals and the conditional parameters in the top row also provide estimates of the effect of the vaccine on infectiousness, none of these parameters produce estimates of the indirect or overall public health effects of vaccination in a population under dependent happenings (5, 6). The effects of a vaccination program could be quite different for people not receiving a vaccine than for those who receive the vaccine. The population-level effects of vaccination depend on the coverage and distribution of the vaccine, as well as on how the different groups mix with each other. These effects could result from the biologic effects as well as the behavioral effects of vaccination. The overall public health effects of a vaccination program depend on the effects in both the vaccinated people and the unvaccinated. Thus, it is important to differentiate among three main types of population-level effects. The *indirect* effects are the population-level effects of widespread vaccination on people not receiving the vaccine. The combination of the population-level effects of widespread vaccination in the individuals receiving the vaccine we call the *total* effects of vaccination and the vaccination program. The *overall* public health effects of the vaccination program depend on the weighted average of indirect effects on the individuals not receiving the intervention and the total effect on the individuals receiving the intervention.

To emphasize the difference between indirect and total effects, we consider a study in which change in average age at infection is an outcome measure of interest. Assume that the vaccine reduces the probability of becoming infected upon exposure to an infectious agent but does not completely protect against infection. Widespread vaccination would reduce the level of transmission and result in an increase in the average age at first infection for both the unvaccinated individuals and the vaccinated individuals. The increase in average age at first infection would be even greater in those who were vaccinated than in those who were not, however, since they would need to be exposed to infection more often, on average, to become infected than those people who did not get vaccinated. Thus, the *indirect* effect of the vaccination program would be less than the *total* effect of the vaccination program. The distinction between these two types of effects is not made often enough. The *overall* change in age of infection would be a weighted average of the increase in the two groups. Even if the overall benefits of vaccination in a population are positive, subgroups of the population may actually suffer; thus, such distinctions must be made.

The indirect, total, and overall effects are defined *within the context of a particular intervention program* in a population. A vaccination program or other kind of intervention program depends on the level of coverage and distribution of the vaccine or other intervention within the population. To estimate the indirect, total, or overall effects of vaccination, we need to compare different populations or communities with and without the intervention program, or with differing types of intervention programs. The three

different kinds of population-level effects—indirect, total, and overall—motivate three types of study designs depending on the choice of comparison populations. In the expression VE = 1 − RR, the estimate of relative risk is derived from comparing different populations. Together with the study designs for individual effects presented above, these are the four study designs for dependent happenings (5, 6) (figure 2). In table 1, the study designs for population-level effects are shown in the right three columns and include study designs IIA, IIB, and III from figure 2. The parameters discussed above are study designs of type I, shown in the third column of the table.

Since the population-level effects of interest are defined within the context of a particular intervention program, the unit of inference is the population, and several populations might be included in the study. The issues raised are similar to those in community trials of interventions against noninfectious diseases. It is important in designing such studies that the populations or communities chosen as the unit of observation be separated as much as possible in every way that is relevant for transmission of the infection under study. The separation could be geographic, cultural, or temporal. We could compare neighboring villages, different cities, or child-care centers with different levels and distributions of vaccination. If there is transmission-relevant mixing among the population units, this could bias the estimates of effectiveness. One population could be the preintervention population while the other population was the postintervention population, raising problems of historical controls. Studies comparing interventions in different populations can be done either as part of phase III trials or as part of postlicensure surveillance. Increasing interest is being shown in evaluation of potential indirect and overall effects of vaccination strategies before licensure.

To illustrate the choice of comparison populations or subpopulations for estimation of the three different effects, assume for simplicity as in figure 2 that there are two populations, A and B. Assume that a vaccination program is implemented in population A so that some, but not necessarily all, of the individuals are vaccinated. No one is vaccinated in population B. To estimate the indirect effects of the vaccination program, we compare the average outcome in the individuals who did not receive the vaccine in the vaccinated population A with the average outcome in unvaccinated individuals in population B. These are study designs of type IIA, represented by the fourth column in table 1. For estimation of the total effect of vaccination, the average outcome in vaccinated people in population A is compared with the average outcome of unvaccinated people in population B. This is represented in the fifth column of table 1. Somewhat more formal definitions of the indirect and total effects are provided elsewhere (4). The effect parameters are distinguished by the subscript specifying the study design and the measure of relative risk. For example, an estimate of the indirect effect of vaccination based on the cumulative incidence at some time T is denoted $VE_{IIA,CI}(T)$.

Subpopulations of the communities could also be compared. For example, the populations could be further stratified by age, sex, receipt of behavioral interventions, or any other risk factor of interest. For a valid comparison, the analogous subpopulation from each community must be included.

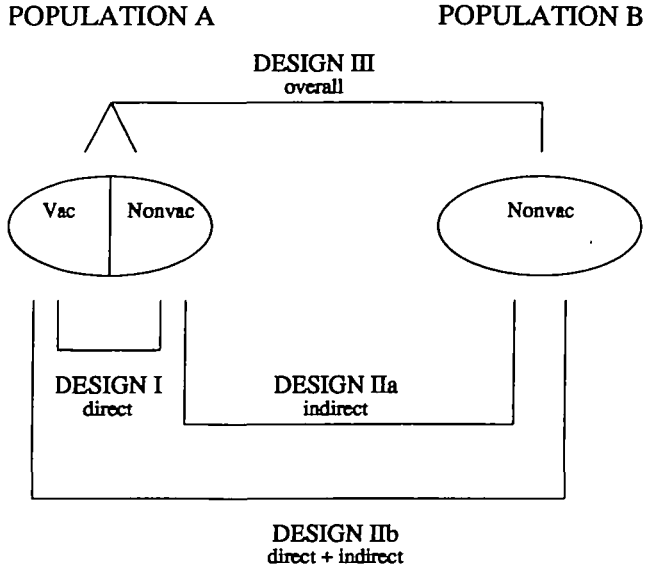

FIGURE 2. Types of effects of vaccination and different study designs for their evaluation based on choice of comparison populations. Populations A and B are separated in every way relevant to transmission dynamics. In population A, some but not necessarily all of the people are vaccinated. In population B, no one is vaccinated. (Adapted from Halloran and Struchiner (6)).

Example

As an example, consider using the hazard rate ratio as the measure of the indirect and total effects of the vaccination program. Let $A0$, $A1$, and $B0$ denote the unvaccinated and vaccinated people in population A and the unvaccinated people in population B, respectively, and let c, p, and $P(t)$ be the contact rate, transmission probability, and prevalence, as defined above. Suppose that the estimated hazard rates in the unvaccinated and vaccinated portions of population A at some time t are $\lambda_{A0}(t) = 0.08$ and $\lambda_{A1}(t) = 0.06$, respectively, and the

estimated hazard rate in the unvaccinated population B is $\lambda_{B0}(t) = 0.12$. Then the estimated indirect and total effects of the vaccination program in population A on the incidence rate compared with no program in population B would be:

$$\text{Indirect:VE}_{\Pi A,\lambda}(t) = 1 - \frac{\lambda_{A0}(t)}{\lambda_{B0}(t)} = 1 - \frac{c_{A0}p_{A0}P_{A0}(t)}{c_{B0}p_{B0}P_{B0}(t)} = 1 - \frac{0.08}{0.12} = 0.33.$$

$$\text{Total:VE}_{\Pi B,\lambda}(t) = 1 - \frac{\lambda_{A1}(t)}{\lambda_{B0}(t)} = 1 - \frac{c_{A1}p_{A1}P_{A1}(t)}{c_{B0}p_{B0}P_{B0}(t)} = 1 - \frac{0.06}{0.12} = 0.50.$$

These expressions show that the indirect or total effects result not only from the change in the transmission probability induced by the direct biologic effects of the vaccination but also from the change induced by the indirect effects. Vaccination would be expected to reduce the prevalence of infective persons in population A. However, mixing patterns might change because of belief in the protective effect of the vaccine. The prevalence of infection in the contacts of vaccinated people in population A, $P_{A1}(t)$, might be higher than the prevalence of infectives among the contacts of the unvaccinated people in population A, $P_{A0}(t)$. Similarly, the contact rates might change in the people who are vaccinated. Thus, the estimated indirect and total effects based on the unconditional parameters can summarize several different consequences of vaccination beyond the biologic consequences.

Note that conducting a trial that includes several different populations or communities does not preclude studying the efficacy of the vaccines within 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 designed to measure the indirect and overall effects of vaccination with different levels of coverage in each population. Within each population, a comparison can be made of the vaccinated portions of the population with the unvaccinated. If information is gathered within the populations on actual contacts, the effect of the vaccine on infectiousness as well as on 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 parameter of interest and the comparison populations should be chosen to provide the effect measures of interest.

In the preceding example, within population A, we could have been conducting a randomized phase III vaccine trial and used the relative hazards in the vaccinated and unvaccinated portions of the population to estimate vaccine efficacy, assuming that the exposure to infection in the two groups remained equal after randomization:

$$\text{VE}_{S,\lambda}(t) = 1 - \frac{\lambda_{A1}(t)}{\lambda_{A0}(t)} = 1 - (0.6/0.8) = 0.25.$$

Although the direct protective efficacy is only 0.25, the combined effect of vaccination and the distribution of vaccinations in the population produces the total protective effect relative to the unvaccinated population of 0.50.

Comparison with prevented fractions in noninfectious diseases

The unconditional vaccine effect parameters discussed above are analogous to the family of prevented fraction parameters discussed by Greenland and Robins (1, 2), with some essential differences. In the work of Greenland and Robins, the prevented fraction in the exposed is estimated by comparing the cumulative incidence in the exposed individuals with the cumulative incidence in the unexposed individuals in a study design similar to study design I with level IV information. The unexposed group is supposed to represent what would have happened to the exposed group had it not been exposed. The number of prevented cases in the exposed can be estimated from the prevented fraction based on the cumulative incidence, if it is known how many people were exposed. A similar argument can be made for the prevented hazard fraction, though the number of cases prevented is not so easily estimated. Under dependent happenings, however, in study design I, in population A, for example, 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 been implemented. If vaccination is widespread enough, the cumulative incidence in the unvaccinated group will usually be lower in the presence of the vaccination program than

it would be 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 this figure is calculated using methods designed for noninfectious diseases, as in study design I and in Greenland and Robins (1, 2).

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 been in place, as in study design IIB. A similar argument applies to estimation of the prevented hazard fraction. Thus, the conceptual framework of the table goes beyond simply extending the family of parameters discussed by Greenland and Robins (1, 2) in two dimensions; it also helps to clarify the similarities and differences between the families of prevented fraction parameters in infectious diseases versus noninfectious diseases.

Reduction in individual infectiousness versus indirect effects

There is an important difference between estimating the reduction in individual infectiousness based on the transmission probability and estimating the indirect effects based on unconditional parameters such as hazard rates or cumulative incidence. The unconditional indirect effect in study design IIA results from the decrease in the exposure to the infectious agent, because 1) fewer people are infected due to reduced transmission and 2) those vaccinated people who become infected might be less infectious. The reduction in infectiousness as estimated using the per-contact reduction in transmission probability does not take into account the lack of infectiousness in people who did not become infected at all because of vaccination. We cannot know who would have become infected without the vaccination program. Thus, with the transmission probability, we obtain a truncated estimate of the reduction in infectiousness, while the indirect effect estimate takes decreased exposure to the infectious agent for a combination of reasons into account.

OVERALL EFFECTIVENESS OF VACCINATION PROGRAMS (VE_{III})

Considering further the study designs for dependent happenings (figure 2), the overall public health benefit of a vaccination program in a population compared with no vaccination is the weighted average of the outcomes in the vaccinated and unvaccinated persons at risk in population A compared with the outcomes in persons at risk in population B. The overall effect depends on the fraction f that is vaccinated in population A. These study designs of type III are represented by the sixth column in table 1. Suppose, in the above example where $\lambda_{A1}(t) = 0.06$ and $\lambda_{A0}(t) = 0.08$, that 75 percent of the people are vaccinated. The overall estimated incidence in population A would be 0.065. The overall public health effectiveness of the vaccination program in reducing the incidence of infection is

$$VE_{III,\lambda}(t) = 1 - \frac{(1-f)\lambda_{A0}(t) + f\lambda_{A1}(t)}{\lambda_{B0}(t)}$$

$$= 1 - (0.065/0.12) = 0.46.$$

The estimated overall effect lies between the indirect (IIA) and total (IIB) effects of the vaccination program. Halloran and Struchiner (6) discuss the study designs for dependent happenings in more detail.

Note that it would be possible to estimate the overall incidence in population A without having to establish who or what fraction had been vaccinated or estimating the indirect or total effects in the vaccinated and unvaccinated subpopulations. *Evaluation* of the overall effects of a vaccination program in a population does not necessarily depend on the mixing patterns within the population, if just the effectiveness in the entire population is of interest. However, the *actual* overall effectiveness of a program will depend heavily on the mixing patterns and the allocation of the vaccine among the different groups. Generally, targeting vaccination to high-transmission groups will have a greater effect. In addition, if one of the measures of interest during a vaccine trial is the overall effectiveness of vaccination, the *design* of the trial could depend on the mixing patterns and the ultimate question of interest. For example, suppose a trial is designed to estimate the efficacy of a vaccine in preschool-age children but the overall effects of the vaccination are also to be evaluated. If most transmission occurs among school-age children and from school-age children to preschool-age children, then, to evaluate the maximum overall effects or to understand the long-term potential overall effects of vaccination, the school-age children might need to be vaccinated even though they are not part of the efficacy trial. If most transmission occurs among preschool-age children, this would not be important. On the other hand, vaccination of either the school-age children or the preschool-age children might reduce the number of events substantially, impairing evaluation of the efficacy. This and many other tradeoffs must be made when designing trials to answer more than one question at a time. Disentangling the direct effects from the indirect effects of vaccines will not always be straightforward.

Example

Monto et al. (30) estimated both the protective efficacy and the overall effect 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 of 1968. The 10-week epidemic period was 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 vaccine efficacy of 0.20 (i.e., $VE_S \approx 1 - (0.072/0.090) = 0.20$), which is rather low. The overall influenza illness attack rate in Tecumseh for the epidemic period was 0.14, while the adjusted overall influenza attack rate 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 schoolchildren is estimated to have been 0.67.

The use of study designs similar to design III, comparing the overall effects of interventions across populations, has been common in infectious disease epidemiology for a long time, but it is becoming more widespread with the integration of more current epidemiologic and statistical methods (31–33). Hayes et al. (34) discuss matching on communities for the randomization and analysis scheme. A design similar to that of study design III is being used currently in a cholera vaccine effectiveness trial in more than 20 communities (J. Clemens, National Institutes of Health, personal communication, 1997) by dividing each community in half geographically and vaccinating all of the individuals in just one half. A similar design will soon be used in community trials of influenza vaccines in matched day-care centers. Evaluation of malaria transmission-blocking vaccines that have only indirect effects in humans will require evaluation comparing separate populations. Many of the lessons from group-randomized designs will be applicable, with special consideration of the dependence of events in infectious diseases. Study designs IIA and IIB, which examine the effects in the vaccinated or unvaccinated subpopulations, have been less common. Hopefully they will become an integral part of vaccine evaluation as investigators become more aware of the differences among the population-level effects in the different groups and combine efficacy and effectiveness evaluation in the same study.

Basic reproductive number (R_0)

Another important parameter of infectious disease epidemiology is the basic reproductive number, R_0. With microparasitic infectious agents, such as viruses, bacteria, and small parasites, the R_0 is the number of new infectives produced by one infectious person in a completely susceptible population during his or her period of infectiousness (35). Similar to the hazard rate, it can be thought of as the product of several components of the transmission system, including the transmission probability, the rate of contact, and the duration of infectiousness. The duration of infectiousness in the expression for R_0 replaces the prevalence of infectives in the expression for the hazard rate. Although it is not included in table 1, R_0 is important in understanding the population dynamics of an infectious agent and the public health effects of intervention. Since a vaccinated person has only a fraction of the transmission probability and possibly a shorter duration of infectiousness, a vaccinated person would not be transmission-dynamically equivalent to a naively susceptible person. The fractional contribution of a vaccinated person to R_0 compared with the contribution of an unvaccinated individual is called the "naive susceptible equivalent" (25). The biologic effect of the vaccine can reduce either the transmission probability per contact or the duration of infectiousness. Thus, estimates of the effect of vaccination on infectiousness, susceptibility, or changes in the duration of infectiousness could provide some information on the effect of vaccination on R_0. For example, if everyone in a population were vaccinated and the reduction in susceptibility were 0.50, the reduction in infectiousness were 0.30, and the reduction in the duration of infectiousness were 0.60, a vaccinated person would be the naive susceptible equivalent of 0.09 of an unvaccinated person. The reproductive number in the vaccinated population would be 0.09 times the original R_0. A vaccine effectiveness measure can also be based on 1 minus the relative basic reproductive numbers in two groups, depending on the choice of comparison populations.

SUMMARY

We have presented a systematic framework of study designs for evaluating different effects of vaccination and vaccination programs depending on the choice of comparison groups and parameter of effect (table 1). We have found this framework useful in organizing our own thoughts and in discussions with other colleagues in designing studies to evaluate vaccination. We hope that others find it useful and that it can contribute to more clarity in the design and analysis of studies of the effects of vaccination on individuals and in populations. Although there are several similarities to the family of prevented risk parameters discussed by Greenland and Robins (1, 2), the dependent hap-

pening structure of infectious diseases adds more levels and types of parameters.

In the third column of table 1, the aim of the studies is to estimate the protective efficacy of vaccination in individuals. This column is divided into two parts, with the amount of information required for estimation increasing from bottom to top (14). This column provides an overview of what is estimable based on the type of data collected in the study. The lower part contains vaccine efficacy based on the unconditional parameters incidence rate ratio ($VE_{S,IR}$), hazard rate ratio ($VE_{S,\lambda}$), and cumulative incidence ratio ($VE_{S,CI}$), which do not require information on actual exposure to infection. These are study designs of type I. If exposure to infection is not equal in the two comparison groups, these designs estimate the combined effects of unequal exposure and the direct biologic protective effects of vaccination. The top row represents the conditional parameters, such as the transmission probability or secondary attack rate. While the third column represents the design needed to estimate the vaccine's effect on susceptibility, $VE_{S,p}$, the fourth and fifth columns represent designs for the effect on infectiousness, $VE_{I,p}$, and the combined effects on the two, $VE_{T,p}$, respectively.

To measure the indirect, total, or overall public health effects of widespread vaccination, it is necessary to compare the effects of vaccination programs in different populations on the parameters of interest. The choice of subpopulations within the different populations determines whether one is measuring the indirect effects, VE_{IIA} (study design IIA), the total effects, VE_{IIB} (study design IIB), or the overall effects VE_{III} (study design III) of the vaccination program. We have emphasized that studies can be designed to answer more than one question at a time.

The fundamental relation between the aspects of exposure to infection and susceptibility to the rate of events clarifies the difference between biologic aspects of the transmission process and the contact aspects. The hierarchy of parameters also provides a framework for thinking about vaccine efficacy studies based on different parameters, especially the conditional versus unconditional parameters, as a missing-data or errors-in-variables problem (4, 7, 36).

We have not discussed vaccine efficacy based on difference parameters here. These were considered by Greenwood and Yule (8) as early as 1915. Difference estimates depend on the absolute incidence in the unvaccinated group, and provide more evidence of the possible public health relevance of vaccines, even when measured within a single population. Table 1 could be entirely rewritten for difference parameters. The parameters listed in table 1 are not exhaustive, but they represent several of the commonly used measures. Many other outcome measures could be used to estimate the effects of a vaccine or a vaccination program. These include average age of first infection and time-to-event as in accelerated failure time models, both of which are related to the hazard rate. Case-control studies are represented by the framework shown here whenever the odds ratio estimator is an approximation of one of the relative risk parameters of table 1.

Increased attention is being given to the design and analysis of vaccine evaluation studies, both pre- and postlicensure. This framework should prove useful in providing a wider and more precise vocabulary for expressing the various kinds of effects being measured and the types of studies needed to estimate the vaccination effects of interest.

ACKNOWLEDGMENTS

This work was partially supported by National Institute of Allergy and Infectious Diseases FIRST Award R29-AI31057 and grant RO1-AI32042. Dr. C. J. Struchiner was partially supported by the Brazilian Research Council (CNPq).

REFERENCES

1. Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fractions. Am J Epidemiol 1988;128:1185–97.
2. Robins JM, Greenland S. Estimability and estimation of excess and etiologic fractions. Stat Med 1989;8:845–59.
3. Ross R. An application of the theory of probabilities to the study of *a priori* pathometry. Part 1. Proc R Soc Ser A 1916;92:204–30.
4. Halloran ME, Struchiner CJ. Causal inference in infectious diseases. Epidemiology 1995;6:142–51.
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. Halloran ME, Struchiner CJ. Study designs for dependent happenings. Epidemiology 1991;2:331–8.
7. Halloran ME, Longini IM Jr, Haber M, et al. Exposure efficacy and change in contact rates in evaluating prophylactic HIV vaccines in the field. Stat Med 1994;13:357–77.
8. Greenwood M, Yule UG. 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.
9. Fox JP, Hall CE, Elveback LR. Epidemiology: man and disease. New York, NY: MacMillan Publishing Company, Inc, 1970.
10. Orenstein WA, Bernier RH, Dondero TJ, et al. Field evaluation of vaccine efficacy. Bull World Health Organ 1985;63:1055–68.
11. Orenstein WA, Bernier RH, Hinman AR. Assessing vaccine efficacy in the field: further observations. Epidemiol Rev 1988;10:212–41.
12. Fine PE, 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.
13. Kendrick P, Eldering G. A study in active immunization against pertussis. Am J Hyg Sect B 1939;38:133.
14. 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.
15. Cox DR. Regression models and life tables (with discussion). J R Stat Soc (B) 1972;34:187–220.
16. 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.
17. Smith PG, Rodrigues LC, Fine PE. Assessment of the protective efficacy of vaccines against common diseases using case-control and cohort studies. Int J Epidemiol 1984;13:87–93.
18. Halloran ME, Haber M, Longini IM Jr. Interpretation and estimation of vaccine efficacy under heterogeneity. Am J Epidemiol 1992;136:328–43.
19. Longini IM Jr, Halloran ME. A frailty mixture model for estimating vaccine efficacy. Appl Stat 1996;45:165–73.
20. Halloran ME, Longini IM Jr, Struchiner CJ. Estimability and interpretation of vaccine efficacy using frailty mixing models. Am J Epidemiol 1996;144:83–97.
21. Struchiner CJ, Halloran ME. Randomization and baseline transmission in vaccine field trials. (Technical report no. 96-7). Atlanta, GA: Emory University, 1996.
22. Durham LK, Longini IM Jr, Halloran ME, et al. Nonparametric estimation of waning vaccine effects using Schoenfeld residuals from the Cox model. (Technical report no. 97-7). Atlanta, GA: Emory University, 1997.
23. Halloran ME, Struchiner CJ, Spielman A. Modeling malaria vaccines. II. Population effects of stage-specific malaria vaccines dependent on natural boosting. Math Biosci 1989;94: 115–49.
24. Halloran ME, Watelet L, Struchiner CJ. Epidemiologic effects of vaccines with complex direct effects in an age-structured population. Math Biosci 1994;121:193–225.
25. Halloran ME, Cochi SL, Lieu TA, et al. Theoretical epidemiologic and morbidity effects of routine varicella immunization of preschool children in the United States. Am J Epidemiol 1994;140:81–104.
26. Koopman JS, Little RJ. Assessing HIV vaccine effects. Am J Epidemiol 1995;142:1113–20.
27. 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.
28. Halloran ME. Evaluating HIV vaccines: discussion. Stat Med 1996;15:2405–12.
29. Longini IM Jr, Sagatelian K, Rida WN, et al. Optimal vaccine trials design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. Stat Med (in press).
30. 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.
31. Grosskurth H, Mosha F, Todd J, et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. Lancet 1995;346: 530–6.
32. Katz J, Carey VJ, Zeger SL, et al. Estimation of design effects and diarrhea clustering within households and villages. Am J Epidemiol 1993;138:994–1006.
33. Clemens J, Brenner R, Rao M, et al. Evaluating new vaccines for developing countries: efficacy or effectiveness? JAMA 1996;275:390–7.
34. Hayes R, Mosha F, Nicoll A, et al. A community trial of the impact of improved sexually transmitted disease treatment on the HIV epidemic in rural Tanzania. 1. Design. AIDS 1995; 9:919–26.
35. Anderson RM, May RM. Infectious diseases of humans: dynamics and control. New York, NY: Oxford University Press, 1991.
36. Golm GT, Halloran ME, Longini IM Jr. Semiparametric models for mismeasured exposure information in vaccine trials. (Technical report no. 97-5). Atlanta, GA: Emory University, 1997.