Published in final edited form as:

Epidemiology. 2011 September ; 22(5): 686–693. doi:10.1097/EDE.0b013e31822708d5.

Bounding the Infectiousness Effect in Vaccine Trials

Tyler J. VanderWeele and Eric J. Tchetgen Tchetgen

Departments of Epidemiology and Biostatistics, Harvard School of Public Health, Boston, MA

Abstract

In vaccine trials, the vaccination of one person might prevent the infection of another; a distinction can be drawn between the ways such a protective effect might arise. Consider a setting with 2 persons per household in which one of the 2 is vaccinated. Vaccinating the first person may protect the second person by preventing the first from being infected and passing the infection on to the second. Alternatively, vaccinating the first person may protect the second by rendering the infection less contagious even if the first is infected. This latter mechanism is sometimes referred to as an “infectiousness effect” of the vaccine. Crude estimators for the infectiousness effect will be subject to selection bias due to stratification on a postvaccination event, namely the infection status of the first person. We use theory concerning causal inference under interference along with a principal-stratification framework to show that, although the crude estimator is biased, it is, under plausible assumptions, conservative for what one might define as a causal infectiousness effect. This applies to bias from selection due to the persons in the comparison, and also to selection due to pathogen virulence. We illustrate our results with an example from the literature.

The vaccine status of one person may affect whether another person is infected. In the statistics literature, this phenomenon is sometimes referred to as “interference”^{1–6}; in the social science literature, as a “spillover effect”^{7,8} or “social interactions”^{9,10}; in the infectious disease context, as an “indirect effect.”^{4,11,12} In the presence of such interference or indirect effects, a further distinction is sometimes drawn in the infectious disease literature. Suppose we consider households of married couples in which one of the 2 individuals is randomized to receive a vaccine or control and the other receives nothing. The vaccine given to one may prevent the other from contracting the disease either because (i) the vaccine prevents the first person from being infected (which then leaves the second person unexposed to the disease) or because (ii) even if the first person is infected, the vaccine may render the infection less contagious, which may in turn prevent the second person from being infected. The latter is sometimes referred to as an “infectiousness effect.”¹³

The infectiousness effect is sometimes estimated by comparing the disease status of the second person in the groups in which the first was vaccinated and infected, versus in the groups in which the first person was unvaccinated and infected.¹⁴ Although intuitively appealing, this crude estimator is subject to selection bias. Conditioning on the infection status of the first person effectively breaks the randomization.¹⁵ For example, those in the vaccinated group who became infected may be overall a less healthy subpopulation than the

Copyright © 2011 by Lippincott Williams & Wilkins

Correspondence: Tyler J. VanderWeele, Harvard School of Public Health, Departments of Epidemiology and Biostatistics, 677 Huntington Avenue, Boston, MA 02115, tvanderw@hsph.harvard.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

unvaccinated group who became infected. If the persons who got the disease even though they were vaccinated are less healthy, they may also be more likely to be contagious and to pass on the disease. The comparison between the 2 infected subgroups is not fair because by conditioning on a variable that occurs after randomization—namely infection status of the first person—we induce selection bias. In addition to selection bias with respect to persons in the study, we may get selection bias arising from differential virulence of the pathogen when different persons are infected. The pathogen virulence among those vaccinated and infected, may for example, be on average greater than among those unvaccinated and infected. Thus, the crude comparison sometimes employed in the literature will in general be biased for the effect it was intended to measure (the “infectiousness effect”).

Recent ideas from causal inference under interference can be used to bound the infectiousness effect in the presence of such selection bias. We use counterfactual notation that allows for such interference (spillover effects) to give a formal counterfactual definition for the “infectiousness effect.” In general, this counterfactual contrast taken as the infectiousness effect will not be identified due to the selection bias described above. However, under some fairly reasonable assumptions, the crude estimator sometimes used to measure the infectiousness effect in fact serves as a lower bound for the formal infectiousness effect causal contrast. The work most closely related to that presented here is the methodology proposed by Hudgens and Halloran.¹⁵ However, Hudgens and Halloran used selection models to identify and conduct sensitivity analysis for the infectiousness effect. Here, we consider assumptions under which the crude estimator will be conservative. Hudgens and Halloran moreover do not consider selection due to pathogen virulence. Finally, in this paper, by more explicitly taking into account interference, we can consider a wider range of effects than do Hudgens and Halloran.

After introducing relevant concepts, we will first consider selection relevant to various persons in the crude comparison, and then consider selection bias arising from varying pathogen virulence. In both cases, we show under reasonable assumptions that this selection bias renders the crude estimator a conservative measure of the true infectiousness effect. We illustrate the results with an example from the vaccine trial literature.¹⁴

Concepts and Definitions

We consider a setting in which there are N households indexed by $i = 1, \dots, N$ in which each household consists of 2 persons indexed by $j = 1, 2$. Let A_{ij} denote the vaccine status for individual j in household i . Let $A_{ij} = 1$ denote that the individual received the vaccination and $A_{ij} = 0$ that the individual did not. For each household, let $A_i = (A_{i1}, A_{i2})$ denote the vaccine status of the 2 individuals in the household. Let Y_{ij} denote the infection status of individual j in household i after some suitable follow-up in the study. Let $Y_{ij}(a_{i1}, a_{i2})$ denote the counterfactual outcome for individual j in household i if the 2 individuals in that household i had, possibly contrary to fact, vaccine status of (a_{i1}, a_{i2}) . For example, $Y_{i2}(1, 0)$ would denote what would have happened to individual 2 if individual 1 had received the vaccine and individual 2 had not; and $Y_{i1}(0, 0)$ denotes what would have happened to individual 1 if neither individual 1 nor individual 2 had received the vaccine. Note that under this counterfactual or “potential outcomes” notation, the potential outcome for individual 1, $Y_{i1}(a_{i1}, a_{i2})$ depends on the vaccine status of both individual 1 and individual 2. This allows for the possibility that the exposure status of one person affects the outcome of another, sometimes referred to as interference or spillover effect. Most literature in causal inference makes a “no interference” assumption^{1,16} that one person's outcome does not depend on the exposure of other people. In the current context, this would imply that $Y_{i1}(a_{i1}, a_{i2}) = Y_{i1}(a_{i1})$ and $Y_{i1}(a_{i1}, a_{i2}) = Y_{i2}(a_{i2})$, so that each person's outcome depends only on his or her own exposure status. This type of no-interference assumption is implausible in the

infectious disease context, and so we do not make it here. We do, however, assume that the exposure status of persons in one household in the study do not affect the outcomes of those in other study households; this is sometimes referred to as an assumption of “partial inference.”^{4,8} This might be plausible if the various households are sufficiently geographically separated or do not interact with one another. Throughout most of this paper, we will assume the simple, randomized experiment described above in which one of the 2 persons is randomized to receive a vaccine or control and the second person is always unvaccinated. We will let $j=1$ denote the individual who may or may not be vaccinated and $j=2$ the individual who is always unvaccinated. The choice of which person is individual 1 and which is individual 2 could be either determined randomly or fixed in advance (eg, in households of married couples, the husband could be the person who is never vaccinated). Because individual 2 is always unvaccinated, we are implicitly conditioning on $A_2 = 0$ throughout. In the penultimate section, we will briefly consider other settings in which in some households both people are vaccinated.

Let us now consider how one might define the “infectiousness effect.” Suppose we are in the setting of a vaccine trial in which in each household i individual 1 is randomized to vaccine or control and individual 2 always receives control. Thus, in half the households, one of the 2 individuals will be vaccinated, and in half the households neither person will be vaccinated. The crude estimator for the infectiousness effect (on the risk difference scale) might be taken as:

$$E[Y_{i2} | A_{i1}=1] - E[Y_{i2} | A_{i1}=0, Y_{i1}=1]. \quad (1)$$

This is a comparison of the infection rates for individual 2 in the subgroup in which individual 1 was vaccinated and infected versus in the subgroup in which individual 1 was unvaccinated and infected. Although this is an appealing intuitive measure for trying to capture the extent to which the vaccine may render those infected less contagious, which may in turn prevent the second individual from being infected (ie, the “infectiousness effect”), the measure is subject to selection bias, as discussed earlier. Although the vaccine status for individual 1, A_{i1} , is randomized, conditioning on a variable that occurs after treatment, (namely, the infection status of individual 1) in effect breaks randomization. As noted earlier, the subgroup with individual 1 vaccinated and infected may be quite different than that in which individual 1 is unvaccinated and infected. We are computing infection rates for individual 2 for subpopulations that are quite different with respect to individual 1. Let us instead consider a second contrast^{6,15}:

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0)=Y_{i1}(0,0)=1]. \quad (2)$$

This contrast compares the infection status for individual 2 if individual 1 was vaccinated, $Y_{i2}(1,0)$ versus unvaccinated, $Y_{i2}(0,0)$, but only among the subset of households for whom individual 1 would have been infected irrespective of whether or not individual 1 was vaccinated ie, $Y_{i1}(1,0) = Y_{i1}(0,0) = 1$. Such a subgroup is sometimes referred to as a principal stratum.^{15,17} Because we are considering only the subset of households for whom individual 1 would have been infected irrespective of whether or not individual 1 was vaccinated, individual 2 is exposed to the infection of individual 1, and thus any effect of the vaccine ought to occur through changing the infectiousness. We might therefore take the contrast in (2) as a formal causal contrast more closely corresponding to the “infectiousness effect.” Moreover, unlike with the crude comparison in (1) we are now comparing, in contrast (2) the infection rates for individual 2 for the same subpopulation, namely the subpopulation for which individual 1 would have been infected irrespective of whether

individual 1 was vaccinated. We are no longer considering a more healthy or less healthy subgroup for individual 1. Note that our framework and definitions do not assume that the second person cannot also be exposed outside the household.

Unfortunately, however, with regard to this causal infectiousness effect we do not know which households fall into this subpopulation in which individual 1 would have been infected irrespective of whether individual 1 was vaccinated. This is because in each household we can only observe the outcome of individual 1 either with the vaccine or without—but not under both scenarios. Because we do not know which households fall into this subpopulation, we cannot compute the contrast in (2) in any straightforward manner from the data. The contrast is, in general, unidentified. In the next section, however, we will show that, under some fairly reasonable assumptions, the crude estimator in (1) is in fact conservative for the causal “infectiousness effect” in contrast (2).

Bounding the Infectiousness Effect

To show that the crude contrast in (1), which can be estimated from data in a trial, is conservative for the causal “infectiousness effect” contrast in (2), we will need 2 assumptions. The first assumption states that the vaccine will never be the cause of the infection; ie, there may be persons who would be infected irrespective of vaccination status or who would not be infected irrespective of vaccination status or who would be infected if unvaccinated and not infected if vaccinated, but there is no one who would be infected if vaccinated and uninfected if unvaccinated. This is stated formally in terms of counterfactual outcomes below.

$$\text{Assumption 1. For all } i, Y_{i1}(1, 0) \leq Y_{i1}(0, 0)$$

Assumption 1 is sometimes referred to as a monotonicity assumption. To show that the crude estimator in (1) is conservative for the causal effect in (2) we need one further assumption. We will state the assumption formally, and then provide some explanation and intuition.

$$\text{Assumption 2. } E[Y_{i2}(0, 0) | A_{i1}=0, Y_{i1}=1] \leq E[Y_{i2}(0, 0) | A_{i1}=1, Y_{i1}=1].$$

Assumption 2 states that the average infection rate for individual 2 if both individuals 1 and 2 were unvaccinated would be lower in the subgroup of households for which individual 1 would be infected and unvaccinated than in the subgroup of households for which individual 1 would be infected and vaccinated. Note that assumption 2 compares 2 subgroups: (i) the subgroup of households for which individual 1 was actually unvaccinated and infected and (ii) the subgroup of households for which individual 1 was actually vaccinated and infected. It then states that if instead we had, contrary to fact, vaccinated no one, then infection rates for individual 2 would be at least as high in the second subgroup as in the first. The assumption is arguably reasonable insofar as the subgroup for which individual 1 was vaccinated and infected is likely less healthy (or the infection more virulent) than the subgroup for which individual 1 was unvaccinated and infected; thus, under the scenario in which both people are unvaccinated, individual 2 is more likely to be infected in the first subgroup than in the second. If this is indeed the case then assumption 2 will hold.

If assumptions 1 and 2 hold then the crude estimator is conservative for the causal infectiousness effect as stated in the following result. The proof of this result and all other are given in the Appendix and eAppendix (<http://links.lww.com/EDE/A493>).

Result 1. Under assumptions 1 and 2, the crude contrast in (1) is conservative for the causal infectiousness effect in (2) in that

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] \leq E[Y_{i2} | A_{i1} = 1, Y_{i1} = 1] - E[Y_{i2} | A_{i1} = 0, Y_{i1} = 1].$$

If in the crude comparison in (1) we find a negative (ie, protective) effect, then the true causal infectiousness effect is even larger in magnitude. If an investigator uses the crude contrast and finds the vaccine of the first person protective for the second person in the subset for whom the first person is infected, then this gives evidence for a true causal infectiousness effective. If the crude estimate is not found protective, this may indicate the absence of a true causal infectiousness effect, or it may be the case that there is a true causal infectiousness effect but the crude estimator, being conservative, is unable to detect it.

Other Measures of Effect

In the previous section, we defined the causal infectiousness effect on a risk difference scale. Other measures of effect might also be of interest. For notational convenience in this section, we define the following:

$$p_v = E[Y_{i2}(1,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1]$$

$$p_u = E[Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1]$$

$$p_1 = E[Y_{i2} | A_{i1} = 1, Y_{i1} = 1]$$

$$p_0 = E[Y_{i2} | A_{i1} = 0, Y_{i1} = 1].$$

The causal infectiousness effect on the risk difference scale is then just $p_v - p_u$ and Result 1 simply states that $p_v - p_u \geq p_1 - p_0$. We might similarly be interested in the causal infectiousness effect on the risk ratio scale, p_v/p_u or on the odds ratio scale, $p_v(1 - p_u)/\{p_u(1 - p_v)\}$. We might further be interested in what one might refer to as the causal vaccine efficacy infectiousness effect, $1 - p_v/p_u$.¹⁵ Result 2 below states that under assumptions 1 and 2, for each of these additional measures of effect, the crude estimator is conservative for the true causal infectiousness effect.

Result 2. Under assumptions 1 and 2,

$$p_v/p_u \leq p_1/p_0,$$

$$p_v(1 - p_u)/\{p_u(1 - p_v)\} \leq p_1(1 - p_0)/\{p_0(1 - p_1)\},$$

$$\text{and } 1 - p_v/p_u \geq 1 - p_1/p_0.$$

Once again, as with result 1, if an investigator found, using the crude estimator, a protective effect on the risk ratio, odds ratio, or vaccine efficacy scales this would be conservative for the true causal effect; the true causal effective would be even more protective than indicated by the crude estimator.

Illustration

Millar et al¹⁴ analyzed results from a group randomized trial in which 7-valent pneumococcal conjugate vaccine (PCV7) was compared with meningococcal conjugate vaccine against serogroup C (MCC) among southwestern American Indian communities. In each household, a child in the household was vaccinated with either the PCV7 vaccine or the MCC vaccine. The primary purpose of the study was to examine whether pneumococcal colonization rates were lower for unvaccinated adults and children in households in which the child was vaccinated with PCV7 versus MCC. The study found a protective effect, with odds ratios of 0.57 (95% CI = 0.33–0.99) and 0.57 (95% CI = 0.26–0.98) for unvaccinated adults and children, respectively.

In addition, the investigators conducted a secondary analysis in which the outcome Y was vaccine type (VT) pneumococcal colonization. They compared the odds of VT colonization for unvaccinated adults for PCV7 versus MCC among households in which the vaccinated child was colonized with a VT strain and obtained an odds ratio of 0.34 (95% CI = 0.11–0.99). This estimate is likely biased for the causal infectiousness effect due to selection and stratifying on a postrandomization variable, colonization status of the vaccinated child. However, under the assumption that the PCV7 vaccine never causes VT pneumococcal colonization (assumption 1) and that VT colonization rates for unvaccinated individuals in households in which the child received the PCV7 vaccine and was colonized would be higher than those in households in which the child received the MCC vaccine was colonized if for both subgroups the child had been given the MCC vaccine (assumption 2, ie, $E[Y_{2L}(0,0) | A_{jL} = 0, Y_{jL} = 1] \geq E[(0,0) | A_{jL} = Y_{jL} = 1]$), then we could conclude that the odds ratio of 0.34 (95% CI = 0.11–0.99) was in fact conservative for the causal “infectiousness” effect odds ratio. We would have evidence for a true infectiousness effect. Note that here the 7-valent vaccine and the colonization status are for the collection of the 7 pneumococcal vaccine types; the assumptions would have to hold for colonization rates of the 7 types taken as a collection.

We also note that the actual design of the Millar et al study was a group randomized trial with 2 to 4 American Indian chapters constituting a randomization unit. Because of this, the “partial interference” assumption that the vaccination status of one household in the study does not affect members of other households will likely be partially violated. The example here is given for illustrative purposes.

Selection Bias Due To Competing Pathogen Strength

Consider now a scenario in which, within a particular household, individual 1 is first exposed to a weak strain of the pathogen such that individual 1 will be infected if unvaccinated but will not be infected if vaccinated; suppose furthermore that later individual

1 is exposed to a more virulent strain of the pathogen such that if individual 1 is not already infected, individual 1 becomes infected by this more virulent strain irrespective of vaccination status. Thus, irrespective of vaccination status, individual 1 is eventually infected (ie, $Y_{i1}(1,0) = Y_{i1}(0,0) = 1$) either by the weaker strain (if unvaccinated) or by the more virulent strain (if vaccinated). The causal infectiousness effect was defined above as: $E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1]$. In the scenario we have just considered, this causal contrast may not actually be capturing the effect of the vaccine on individual 2 by making individual 1 less infectiousness due to selection bias of the following form. Suppose that in the household we have just considered, individual 1, if unvaccinated is infected by the weaker strain and because it is weaker, individual 2 is not infected; suppose further if individual 1 is infected by the more virulent strain when vaccinated, then individual 2 will be infected because of the more virulent strain. In this scenario, it would appear as though the vaccine has a “causative” (ie, detrimental) infectiousness effect, as individual 2 would be infected only if individual 1 is vaccinated. This, however, would be the result of selection, rather than the infectiousness effect of the vaccine. One way to attempt to deal with this would be to consider the strata for which individual 1 is infected irrespective of vaccination status, and then to restrict inference further so that the subset of persons for whom the virulence of the pathogen that would have infected individual 1 would be the same irrespective of vaccination status. For example, among those infected irrespective of vaccination status, we could define $S_i(0)$ and $S_i(1)$ to be the virulence of the pathogen causing the infection when individual 1 was unvaccinated or vaccinated, respectively. We could then consider the infectiousness effect for those with $S_i(0) = S_i(1)$. We could, for different levels of pathogen virulence s , consider

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1, S_i(0) = S_i(1) = s].$$

Unfortunately, even if we have data on the levels of pathogen virulence causing the infection, we will not observe both $S_i(0)$ and $S_i(1)$; we would be able to only observe one of these, $S_i(1)$ if individual 1 is vaccinated and $S_i(0)$ if unvaccinated.

Let S_i denote the virulence of the pathogen that actually caused infection (ie, for those infected, $S_i = S_i(1)$ if individual 1 is vaccinated and $S_i = S_i(0)$ if unvaccinated) and suppose $S_i = 1$ denotes a virulent pathogen and $S_i = 0$ denotes a weaker pathogen. Suppose we had data on the levels of pathogen virulence, S_i , actually causing each infection. Although we cannot identify the infectiousness effect conditional on the principal strata of pathogen virulence $S_i(0) = S_i(1) = s$, we show in the eAppendix (<http://links.lww.com/EDE/A493>) that under assumption 1, a slight modification of assumption 2 (see eAppendix) and a third assumption that $S_i(0) \leq S_i(1)$ (ie, for people who would be infected irrespective of vaccination status, the virulence of the pathogen causing infection when the person is vaccinated is at least as virulent as the pathogen causing the infection when the person is unvaccinated), then the crude estimator

$$E[Y_{i2} | A_{i1}=1, Y_{i1}=1, S_i=0] - E[Y_{i2} | A_{i1}=0, Y_{i1}=1, S_i=0]$$

is conservative (ie, less than the pathogen-virulence conditional infectiousness effect),

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1, S_i(0) = S_i(1) = 0]$$

To show a similar result for the pathogen-virulence conditional infectiousness effect in (3) when conditioning on $S_i(0) = S_i(1) = 1$ rather than $S_i(0) = S_i(1) = 0$, somewhat stronger assumptions are needed. Details are given in the eAppendix (<http://links.lww.com/EDE/A493>).

Other Designs and Effects

Thus far, for simplicity, we have considered a randomized design in which one of 2 persons is randomized to receive a vaccine or control and the other is always unvaccinated. All of the above results would also hold if within each household, one person was randomized to vaccination or control and all others were unvaccinated; in this case, Y_{j2} would simply be a vector. This was in fact the design used in the study of Millar et al.¹⁴

In the study design we have been considering, all households have either one person or no one vaccinated. We might instead consider a design in which each of the 2 persons is randomized to receive vaccine or control, so that in some households neither individual is vaccinated, in some just one, and in some both. As before, if only one person is vaccinated, we will let $j = 1$ denote this individual. In the simpler design in which at most one person is vaccinated, in addition to considering the causal infectiousness effect defined above, we might also consider direct effects of the form $E[Y_{j1}(1,0) - Y_{j1}(0,0)]$ (the effect on individual 1 of individual 1's being vaccinated) or spillover/indirect effects for the form $E[Y_{j2}(1,0) - Y_{j2}(0,0)]$ (the effect on individual 2 of individual 1's being vaccinated). Such effects were considered by Halloran and Struchiner¹¹ and by Hudgens and Halloran.⁴ We note here that in the simple setting considered in this paper, in which each household has only 2 people, these direct and spillover/indirect effects are identified by sample means $E[Y_{j1}|A_{j1} = 1] - E[Y_{j1}|A_{j1} = 0]$ and $E[Y_{j2}|A_{j1} = 1] - E[Y_{j2}|A_{j1} = 0]$ and inference reduces simply to inference for a difference in means. When households have more than 2 people, the formal definitions of these direct and indirect effects become more subtle^{4,18} and statistical inference is rendered considerably more complex—although it is still possible.⁶

When we move to the more general designs in which each of the 2 persons is randomized to receive vaccine or control, we can also define direct effects of the form $E[Y_{j2}(1,1) - Y_{j2}(1,0)]$ (the effect on individual 2 of individual 2's being vaccinated when individual 1 is also vaccinated) or spillover/indirect effects for the form $E[Y_{j1}(1,1) - Y_{j1}(1,0)]$ (the effect on individual 1 of individual 2's being vaccinated when individual 1 is also vaccinated). With just 2 people in each household, these effects are also identified by a difference in sample means $E[Y_{j2}|A_{j1} = 1, A_{j2} = 1] - E[Y_{j2}|A_{j1} = 1, A_{j2} = 0]$ and $E[Y_{j1}|A_{j1} = 1, A_{j2} = 1] - E[Y_{j1}|A_{j1} = 1, A_{j2} = 0]$, respectively, and inference for these effects reduces to inferences for comparing 2 means. Furthermore, in this more general design, in addition to the causal infectiousness defined above in (2) we could also define a further causal infectiousness effect,

$$E[Y_{i1}(1,1) - Y_{i1}(1,0) | Y_{i2}(1,1) = Y_{i2}(1,0) = 1] \quad (4)$$

that is, the effect on individual 1 of individual 2's receiving the vaccine when individual 1 is also vaccinated within the subpopulation for whom individual 2 would be infected irrespective of whether individual 2 received the vaccine. As described in the Appendix under this more general design the crude estimator, $E[Y_{j1}|A_{j1} = 1, A_{j2} = 1, Y_{j2} = 1] - E[Y_{j1}|A_{j1} = 1, A_{j2} = 0, Y_{j2} = 1]$, will be conservative for this causal infectiousness effect under assumptions similar to assumptions 1 and 2 above.

Discussion

We have used causal inference theory concerning interference and principal stratification to provide bounds on what can be defined as an infectiousness effect. We have shown under fairly plausible assumptions that the crude estimator for the infectiousness effect will be conservative for what can be defined as a true causal infectiousness effect; this holds for a variety of effect measures. The selection biases from conditioning on a postvaccination variable due to either health status or pathogen virulence biases the estimate in a conservative manner. The results will be useful in reasoning about the possible presence of an infectiousness effect in vaccine trials. The results given here could potentially be extended to longitudinal settings and to other types of study designs.

In a sense, our paper constitutes “good news” for the field of infectious disease research concerning randomized trials of the type considered above insofar as, if in these studies a significant infectiousness effect is found using the crude measure, then researchers can be reasonably confident that the true infectiousness effect is of even larger magnitude. On the other hand, it is possible that in studies in which the infectiousness effect of the vaccine was not found significant, this may have been due simply to the downward bias due to selection and an infectiousness effect may still in fact be present.

The results given here are of course limited by the assumptions made. We have focused on 2 assumptions. The first was that the vaccine never causes infection. The second assumption compared 2 subgroups: (i) the subgroup of households for which individual 1 was actually unvaccinated and infected and (ii) the subgroup of clusters for which individual 1 was actually vaccinated and infected; the second assumption then required that if we had, contrary to fact, vaccinated no one, then infection rates for individual 2 would be at least as high in the second subgroup as in the first. We believe that in many settings, these assumptions will be plausible. However, whether the assumptions hold will depend on the nature of the vaccine under study; the assumptions will not be applicable to all vaccines. The framework used here also made an assumption of “partial inference” that the exposure status of persons in one household in the study do not affect the outcomes of those in other households in the study. This might be plausible if the various households are sufficiently geographically separated or do not interact with one another. Importantly, this assumption pertains to household units in the study, not to all households that might have been in the study. In certain settings, this assumption of partial interference might be plausible.

The assumption has implications for study design in 2 important and contrasting respects. On the one hand, the assumption will be more plausible if the households selected for the study are geographically separated with relatively little chance of the study participants interacting with one another; this might motivate the selection of households that are geographically separated. On the other hand, in the context of a group randomized trial, it might be hoped that if groups of households interacting within one another are all vaccinated, this will lead to further protective “indirect” or “spillover” effects; this would motivate selecting households that are in geographic proximity to one another. This was in fact the study design used by Millar et al,¹⁴ and for precisely this reason. Future work could attempt to generalize the results here to allow for violations of the partial interference assumption. Importantly, however, one setting in which the current results in this paper would be applicable is that of an individually randomized trial in which outcome data are also collected on households, with the study population from which the study sample is drawn being large enough so that households in the study sample do not interact with one another.

We have focused on the setting of a randomized trial, but our results are potentially applicable to observational studies as well. Our results would hold in an observational study if, conditional on some set of covariates C , the treatment was jointly independent of the counterfactual outcomes (ie, effectively randomized within strata of C). The assumptions about conditional expectations (assumption 2) would also have to be made conditional on C . That such results hold in observational studies extends considerably the potential applicability of our results. Suppose, for example, that in an observational study the treatment were instead a smoking cessation program in which one of 2 persons in a household participated. The participation of the first person might affect the smoking behavior of the second. This might occur either (i) because smoking cessation for the first person encourages the second to stop smoking or because (ii) even if the first person does not stop smoking, the second person might nevertheless be exposed to some of the smoking cessation program materials. One could potentially test for the presence of this second type of effect by applying the results we have given above concerning the “infectiousness effect.” The results given in the paper could likewise be applicable to a range of health-related, social and psychologic outcomes and exposures. The methodology described here may facilitate drawing inferences in such settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Elizabeth Halloran, 2 referees, and the editor for helpful comments.

Supported by National Institutes of Health (grant HD060696).

Appendix 1

Proofs of Results 1 and 2

Proof of Result 1

Under assumption 1, we have that

$$\begin{aligned} & E[Y_{i2}(1,0) | Y_{i1}(1,0)=Y_{i1}(0,0)=1] \\ &= E[Y_{i2}(1,0) | A_{i1}=1, Y_{i1}(1,0)=Y_{i1}(0,0)=1] \\ &= E[Y_{i2}(1,0) | A_{i1}=1, Y_{i1}(1,0)=1] \\ &= E[Y_{i2} | A_{i1}=1, Y_{i1}=1] \end{aligned}$$

where the first equality holds by randomization of A_{i1} , the second by assumption 1 (monotonicity of the vaccine), and the third follows because if $A_{i1}=1$, $A_{i2}=0$ then $Y_{i2}(1,0) = Y_{i2}$ and $Y_{i1}(1,0) = Y_{i1}$ (ie, the “consistency” assumption). Also under assumption 1,

$$\begin{aligned} & E[Y_{i2}(0,0) | Y_{i1}(1,0)=Y_{i1}(0,0)=1] \\ &= E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}(1,0)=Y_{i1}(0,0)=1] \\ &= E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}(1,0)=1] \\ &= E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}=1] + \{E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1] - E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1]\} \\ &= E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}=1] + \{E[Y_{i2} | A_{i1}=0, Y_{i1}=1] - E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1]\} \\ &= E[Y_{i2} | A_{i1}=0, Y_{i1}=1] + \{E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}=1] - E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1]\} \end{aligned}$$

where the first equality holds by randomization of A_{i1} and the second by assumption 1 (monotonicity of the vaccine); the third follows by adding and subtracting, and the fourth because if $A_{i1} = 1, A_{i2} = 0$ then $Y_{i1}(1,0) = Y_{i1}$ and if $A_{i1} = 0, A_{i2} = 0$ then $Y_{i2}(0,0) = Y_{i2}$ (ie, the “consistency” assumption). We thus have:

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] \\ = E[Y_{i2} | A_{i1}=1, Y_{i1}=1] - E[Y_{i2} | A_{i1}=0, Y_{i1}=1] + \{E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1] - E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}=1]\}.$$

Under assumption 2, we then have

$$E[Y_{i2}(1,0) - Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] \leq E[Y_{i2} | A_{i1}=1, Y_{i1}=1] - E[Y_{i2} | A_{i1}=0, Y_{i1}=1].$$

Proof of result 2

It was shown in the Proof of Result 1 that under assumption 1,

$$p_v = E[Y_{i2}(1,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] = E[Y_{i2} | A_{i1}=1, Y_{i1}=1] = p_1$$

and

$$E[Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] \\ = E[Y_{i2} | A_{i1}=1, Y_{i1}=1] + \{E[Y_{i2}(0,0) | A_{i1}=1, Y_{i1}=1] - E[Y_{i2}(0,0) | A_{i1}=0, Y_{i1}=1]\}$$

so that under assumptions 1 and 2 we have

$$p_u = E[Y_{i2}(0,0) | Y_{i1}(1,0) = Y_{i1}(0,0) = 1] \geq E[Y_{i2} | A_{i1}=0, Y_{i1}=1] = p_0.$$

By Result 1, $p_v - p_u \leq p_1 - p_0$. Since $p_u \geq p_0$, we thus also have, by dividing, that $(p_v - p_u)/p_u \leq (p_1 - p_0)/p_0$ ie, $p_v/p_u - 1 \leq p_1/p_0 - 1$. Multiplying this inequality by -1 gives the result for the vaccine efficacy measure. Moreover, from $p_v/p_u - 1 \leq p_1/p_0 - 1$, we immediately have $p_v/p_u \leq p_1/p_0$, which gives the result for the risk ratio measure. Because, under assumption 1, $p_v = p_1$ we thus also have $p_v/\{p_u(1 - p_v)\} \leq p_1/\{p_0(1 - p_v)\}$, and, because, under assumptions 1 and 2, $p_u \geq p_0$ and so $(1 - p_u) \leq (1 - p_0)$, we also have $p_v(1 - p_u)/\{p_u(1 - p_v)\} \leq p_1(1 - p_0)/\{p_0(1 - p_1)\}$, which gives the result for the odds ratio measure.

Appendix 2

Causal Infectiousness Effect for Designs in Which Both Individuals Are Randomized to Receive the Vaccine

If each of the 2 individuals is randomized to receive vaccine or control, then, by an argument similar to that in the Proof of Result 1, the crude estimator, $E[Y_{i2} | A_{i1} = 1, A_{i2} = 0, Y_{i1} = 1] - E[Y_{i2} | A_{i1} = 0, A_{i2} = 0, Y_{i1} = 1]$, will still be conservative for the causal

infectiousness effect in (2) if Assumption 1 holds, along with explicitly conditioning on $A_{i2} = 0$ in Assumption 2, ie,

$$\text{Assumption2}^{\dagger}. E[Y_{i2}(0,0) | A_{i1}=0, A_{i2}=0, Y_{i1}=1] \leq E[Y_{i2}(0,0) | A_{i1}=1, A_{i2}=0, Y_{i1}=1].$$

Moreover, again by an argument similar to that in the Proof of Result 1, the crude estimator, $E[Y_{i1} | A_{i1} = 1, A_{i2} = 1, Y_{i2} = 1] - E[Y_{i1} | A_{i1} = 1, A_{i2} = 0, Y_{i2} = 1]$, is conservative for the causal infectiousness effect in, (4) namely,

$$E[Y_{i1}(1,1) - Y_{i1}(1,0) | Y_{i2}(1,1)=Y_{i2}(1,0)=1]$$

provided that the following 2 assumptions hold:

$$\text{Assumption1}^{\dagger\dagger}. \text{For all } i, Y_{i2}(1,1) \leq Y_{i2}(1,0)$$

$$\text{Assumption2}^{\dagger\dagger}. E[Y_{i1}(1,0) | A_{i1}=1, A_{i2}=0, Y_{i2}=1] \leq E[Y_{i1}(1,0) | A_{i1}=1, A_{i2}=1, Y_{i2}=1].$$

Proof is provided in the eAppendix (<http://links.lww.com/EDE/A493>).

References

1. Cox, DR. The Planning of Experiments. New York: Wiley; 1958.
2. Hong G, Raudenbush SW. Evaluating kindergarten retention policy: a case study of causal inference for multilevel observational data. *J Am Stat Assoc.* 2006; 101:901–910.
3. Rosenbaum PR. Interference between units in randomized experiments. *J Am Stat Assoc.* 2007; 102:191–200.
4. Hudgens MG, Halloran ME. Towards causal inference with interference. *J Am Stat Assoc.* 2008; 103:832–842. [PubMed: 19081744]
5. VanderWeele TJ. Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Sociol Methods Res.* 2010; 38:515–544.
6. Tchetgen Tchetgen EJ, VanderWeele TJ. Estimation of causal effects in the presence of interference [special Issue on Causal Inference]. *Stat Methods Med Res.*
7. Strain PS, Shores RE, Kerr MM. An experimental analysis of “spillover” effects on the social interaction of behaviorally handicapped preschool children. *J Appl Behav Anal.* 1976; 9:31–40. [PubMed: 1254540]
8. Sobel ME. What do randomized studies of housing mobility demonstrate? Causal Inference in the face of interference. *J Am Stat Assoc.* 2006; 101:1398–1407.
9. Graham B. Identifying social interactions through conditional variance restrictions. *Econometrica.* 2008; 76:643–660.
10. Manski, CF. Identification of treatment response with social interactions. Working Paper; 2010; Northwestern University;
11. Halloran ME, Struchiner CJ. Study designs for dependent happenings. *Epidemiology.* 1991; 2:331–338. [PubMed: 1742381]
12. Halloran ME, Struchiner CJ. Causal inference for infectious diseases. *Epidemiology.* 1995; 6:142–151. [PubMed: 7742400]

13. Datta S, Halloran ME, Longini IM. Efficiency of estimating vaccine efficacy for susceptibility and infectiousness: randomization by individual versus household. *Biometrics*. 1999; 55:792–798. [PubMed: 11315008]
14. Millar EV, Watt JP, Bronsdon MA, et al. Indirect effect of 7-valent pneumococcal conjugate vaccine on pneumococcal colonization among unvaccinated household members. *Clin Infect Dis*. 2008; 47:989–996. [PubMed: 18781875]
15. Hudgens MG, Halloran ME. Causal vaccine effects on binary post-infection outcomes. *J Am Stat Assoc*. 2006; 101:51–64. [PubMed: 19096723]
16. Rubin DB. Comment on: Neyman (1923) and Causal Inference in Experiments and. *Observational Studies*. *Stat Sci*. 1990; 5:472–480.
17. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002; 58:21–29. [PubMed: 11890317]
18. VanderWeele TJ, Tchetgen Tchetgen EJ. Effect decomposition under inference for two-stage randomized experiments. In press.