Scientific
Research

A Simple Method of Measuring Vaccine Effects on Infectiousness and Contagion

Yasutaka Chiba

Division of Biostatistics, Clinical Research Center, Kinki University School of Medicine, Osaka, Japan
Email: chibay@med.kindai.ac.jp

ABSTRACT

The vaccination of one person may prevent another from becoming infected, either because the vaccine may prevent the first person from acquiring the infection and thereby reduce the probability of transmission to the second, or because, if the first person is infected, the vaccine may impair the ability of the infectious agent to initiate new infections. The former mechanism is referred as a contagion effect and the latter is referred as an infectiousness effect. By applying a principal stratification approach, the conditional infectiousness effect has been defined, but the contagion effect is not defined using this approach. Recently, new definitions of unconditional infectiousness and contagion effects were provided by applying a mediation analysis approach. In addition, a simple relationship between conditional and unconditional infectiousness effects was found under a number of assumptions. These two infectiousness effects can be assessed by very simple estimation and sensitivity analysis methods under the assumptions. Nevertheless, such simple methods to assess the contagion effect have not been discussed. In this paper, we review the methods of assessing infectiousness effects, and apply them to the inference of the contagion effect. The methods provided here are illustrated with hypothetical vaccine trial data.

1. Introduction

Evaluating the effect of vaccination on reducing infectiousness has important public health consequences [1]. Even if a vaccine does not provide strong protection against an infection, it could substantially reduce the total number of cases if transmission from an infected vaccinated person is reduced compared to that from a non-vaccinated person. This is because the vaccine status of one person may affect whether another person becomes infected. This phenomenon is referred as "interference" in the statistical literature [2], or the "indirect effect" in the infectious disease context [3]. In the presence of such interference or indirect effects, a further distinction is drawn, as mentioned below.

Considering households consisting of two persons, one (person 1) is randomized to receive a vaccine or a control, and the other (person 2) receives nothing. The vaccination of person 1 may prevent person 2 from becoming infected via the following two mechanisms: 1) the vaccine may prevent person 1 from acquiring the infection and thereby reduce the probability of transmission to person 2, or 2) if person 1 is infected irrespective of the

vaccine, the vaccine may impair the ability of the infectious agent to initiate new infections; *i.e.*, make the agent less infectious. The first of these mechanisms is referred as the "contagion effect" [4] and the second is referred as the "infectiousness effect" [5].

To give a formal definition of these two effects, the principal stratification and mediation analysis approaches have been adapted. Recently, by applying the principal stratification approach, the conditional infectiousness effect was defined [3,6], but unfortunately the contagion effect was not. No one may be able to define it using this approach. More recently, unconditional infectiousness and contagion effects were defined by applying the mediation analysis approach [4]. Furthermore, a simple relationship between conditional and unconditional infectiousness effects was found under a number of assumptions [7]. These two infectiousness effects can be assessed using very simple statistical methods under the assumptions. Nevertheless, such methods to assess the contagion effect have not been discussed. In this paper, we review the methods used to assess infectiousness effects, and apply them to the inference of the contagion

effect.

This paper is organized as follows. Section 2 presents the concepts and definitions used throughout this paper. Section 3 reviews the relationship between the conditional and unconditional infectiousness effects, and expresses the contagion effect in terms of principal stratification. In Section 4, we describe a simple method of estimating these effects upon their identification, and provide a simple sensitivity analysis method of assessing how inferences would change under violations of the identification assumption in Section 5. The methods presented in Sections 4 and 5 are illustrated using hypothetical randomized vaccine trial data in Section 6. Section 7 concludes with a discussion.

2. Concepts and Definitions

The notation and fundamental assumptions used throughout this paper are presented in Section 2.1. Section 2.2 presents the crude estimator of the infectiousness effect, and indicates the problem with this estimator. Section 2.3 formalizes the conditional infectiousness effect by applying the principal stratification approach, and in Section 2.4 the unconditional infectiousness and contagion effects are formalized by applying the mediation analysis approach.

2.1. Notation and Assumptions

We consider a setting in which there are N households indexed by $i = 1, \cdots, N$, in which each household consists of two persons indexed by $j = 1, 2$. Let A_{ij} denote the vaccine status of person j in household i, where $A_{ij} = 1$ if the person received the vaccination and $A_{ij} = 0$ if the person did not. Let Y_{ij} denote the infection status of person j in household i, where $Y_{ij} = 1$ if the person was infected and $Y_{ij} = 0$ if the person was not. Finally, let $X_i = (X_{i1}, X_{i2}, \cdots, X_{im})$ denote a set of baseline covariates for household i. Because person 1 is randomized and person 2 receives nothing, X_i affects Y_{i1} and Y_{i2} but does not affect A_{i1} and A_{i2}. The relationship among these variables is represented by a directed acyclic graph (DAG) [8,9] (**Figure 1**), in which A_2 is not displayed because the vaccine status of person 2 (A_2) is identical among all households (*i.e.*, person 2 is never vaccinated) ($A_2 = 0$).

We assume that the vaccine status of the persons in one household does not affect the outcomes of those in other households; this is sometimes referred to as an assumption of partial interference [3,10]. We let $Y_{ij}(a_{i1}, a_{i2})$ denote the potential outcome for person j in household i if the two persons in household i had, possibly contrary to fact, a vaccine status of (a_{i1}, a_{i2}). This assumption of partial interference might be plausible if the various households are sufficiently geographically separated or do not interact. We further require the con-

sistency assumption, which means that the value of Y_{ij} that would have been observed if A_{ij} had been set to what in fact it was is equal to the value of Y_{ij} that was observed; *i.e.*, $Y_{ij} = Y_{ij}(A_{i1}, A_{i2})$ [11]. In addition, by random assignment to person 1, it is assumed that $Y_{ij}(a_{i1}, a_{i2})$ is independent of A_{i1}. This independency can also be assumed conditional on X_i. Because person 2 is always unvaccinated, we simplify the notation as $Y_{ij}(a_{i1}) := Y_{ij}(a_{i1}, 0)$.

Using this notation, on the vaccine efficacy scale, the indirect effect is formalized by

$$1 - \Pr(Y_{i2}(1) = 1) \big/ \Pr(Y_{i2}(0) = 1)$$

(*i.e.*, the effect on person 2 of person 1 being vaccinated) [3]. In contrast, the effect on person 1 of person 1 being vaccinated is called the direct effect, and is formalized by

$$1 - \Pr(Y_{i1}(1) = 1) \big/ \Pr(Y_{i1}(0) = 1).$$

In the current setting, due to random assignment to person 1 and the consistency assumption, these direct and indirect effects are identified by the sample proportions

$$1 - \Pr(Y_{i1} = 1 \mid A_{i1} = 1) \big/ \Pr(Y_{i1} = 1 \mid A_{i1} = 0),$$

$$1 - \Pr(Y_{i2} = 1 \mid A_{i1} = 1) \big/ \Pr(Y_{i2} = 1 \mid A_{i1} = 0),$$

respectively. The indirect effect is classified into infectiousness and contagion effects due to the existence of two infection mechanisms, as noted in Section 1.

2.2. Crude Estimator of the Infectiousness Effect

The crude estimator of the infectiousness effect might be taken as [12]:

$$1 - \frac{\Pr(Y_{i2} = 1 \mid A_{i1} = 1, Y_{i1} = 1)}{\Pr(Y_{i2} = 1 \mid A_{i1} = 0, Y_{i1} = 1)}. \qquad (2.1)$$

This is a comparison of the infection proportions for person 2 in the subgroup in which person 1 was vaccinated and infected *versus* that in which person 1 was unvaccinated and infected. This is an appealing intuitive method of capturing the extent to which the vaccine may render those infected less contagious, which may in turn prevent the second person from being infected (*i.e.*, infectiousness effect).

However, the measure is subject to selection bias. Although the vaccine status of person 1, A_{i1}, is randomized, conditioning on a variable that occurs after treatment (*i.e.*, the infection status of person 1) breaks the randomization. This can be indicated using the DAG shown in **Figure 1**. Even if an arrow from X to A_1 does not exist due to randomization, conditioning for Y_1 would induce non-identifiability due to the induced structural relationship of a possible double arrow between A_1 and X, as shown in **Figure 2** [13].

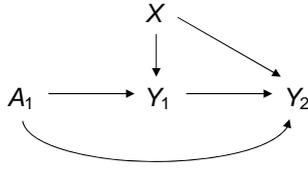As a result, the subgroup with person 1, who was vac-

Figure 1. A directed acyclic graph representing the relationship among A_1, Y_1, Y_2, and X in the current context.
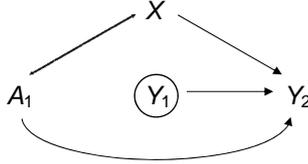

Figure 2. A graph after conditioning on Y_1.

cinated and infected, may be quite different from that in which person 1 is unvaccinated and infected. For example, those in the vaccinated group who become infected may be a less healthy subpopulation than those in the unvaccinated group who become infected. If the persons who are infected are less healthy even though they have been vaccinated, they may also more likely to be contagious and to pass on the disease. We are then computing infection proportions for person 2 for subpopulations that are quite different with respect to person 1.

2.3. Conditional Infectiousness Effect

By applying the principal stratification approach, consider the following contrast [14]:

$$I_C := 1 - \frac{\Pr\left(Y_{i2}(1)=1 \mid Y_{i1}(1)=Y_{i1}(0)=1\right)}{\Pr\left(Y_{i2}(0)=1 \mid Y_{i1}(1)=Y_{i1}(0)=1\right)}. \qquad (2.2)$$

This compares the infection status of person 2 if person 1 was vaccinated, $Y_{i2}(1)$, *versus* that if person 1 was unvaccinated, $Y_{i2}(0)$, but only in the subgroup of households in which person 1 would have been infected irrespective of whether person 1 was vaccinated; *i.e.*, $Y_{i1}(1)=Y_{i1}(0)=1$. Such a subgroup is sometimes referred to as a principal stratum [15]. Because we are considering only the subgroup of households for whom person 1 would have been infected irrespective of whether person 1 was vaccinated, person 2 is exposed to the infection of person 1, and thus any effect of the vaccine ought to occur through a change in the infectiousness. Therefore, the conditional infectiousness effect is defined by Equation (2.2), because it is conditioned on $Y_{i1}(1)=Y_{i1}(0)=1$. Moreover, unlike with the crude comparison in Equation (2.1), we are now comparing the infection proportions of person 2 for the same subpopulation in Equation (2.2). We are no longer considering a healthier or unhealthier subgroup for person 1.

Unfortunately, we do not know which households fall

into the subpopulation in which person 1 would have been infected irrespective of whether he or she was vaccinated. This is because we can observe only the outcome of person 1, either with or without the vaccine but not under both scenarios. Because we do not know which households fall into this subpopulation, we cannot compute the conditional infectiousness effect in a straightforward manner. However, in Section 4, we will show that the effect can be estimated simply under a number of assumptions.

2.4. Unconditional Infectiousness and Contagion Effect

Suppose that, in addition to potentially intervening to vaccinate person 1, we could, at least hypothetically, intervene to infect or not infect person 1. Then, $Y_{i2}(a_{i1}, a_{i2}, y_{i1})$ would denote the infection status of person 2 if we would set the vaccine status of person 1 and person 2 to a_{i1} and a_{i2} and the infection status of person 1 to y_{i1}. The assumption that person 2 is always unvaccinated allows a simplified notation: $Y_{i2}(a_{i1}, y_{i1}) := Y_{i2}(a_{i1}, 0, y_{i1})$. This potential outcome is used to define unconditional infectiousness and contagion effects [4].

Consider the following contrast:

$$I_U := 1 - \frac{\Pr\left(Y_{i2}(1, Y_{i1}(1))=1\right)}{\Pr\left(Y_{i2}(0, Y_{i1}(1))=1\right)}. \qquad (2.3)$$

This compares the potential infection status of person 2 if person 1 had been vaccinated *versus* unvaccinated and person 1 had the infection status that would occur if vaccinated. If Equation (2.3) is non-zero, this will be because even when person 1 is vaccinated and infected, the vaccine itself affects whether person 2 is infected by person 1. In some ways, it is analogous to what is called an infectiousness effect. However, this measure defined by Equation (2.3) differs from the conditional infectiousness effect defined by Equation (2.2) in that it is not conditional on person 1 actually being infected.

Consider now the other contrast:

$$C := 1 - \frac{\Pr\left(Y_{i2}(0, Y_{i1}(1))=1\right)}{\Pr\left(Y_{i2}(0, Y_{i1}(0))=1\right)}. \qquad (2.4)$$

The term $Y_{i2}(0, Y_{i1}(1))$ considers what the potential infection status of person 2 is if person 1 is unvaccinated, but we set the infection status of person 1 to the level it would have been if person 1 was vaccinated. Equation (2.4) compares this potential outcome to $Y_{i2}(0, Y_{i1}(0))$, which is the potential infection status of person 2 if person 1 is not vaccinated, and we set the infection status of person 1 to the level it would be if person 1 was unvaccinated. For Equation (2.4) to be nonzero, $Y_{i1}(1)$ and

OJS

$Y_{i1}(0)$ have to differ; *i.e.*, vaccination of person 1 would have to affect the infection status of person 1, and that change in infection for person 1 would have to change the infection status for person 2, even if person 1 had remained unvaccinated. Essentially, Equation (2.4) is non-zero if the vaccine prevents infection in person 1,

and that in turn prevents person 2 from being infected. Thus, the contagion effect is defined by Equation (2.4).

These definitions of the unconditional infectiousness and contagion effects have the feature that we can decompose an indirect effect into unconditional infectiousness and contagion effects; *i.e.*, from Equations (2.3) and (2.4),

$$1 - \frac{\Pr\left(Y_{i2}(1)=1\right)}{\Pr\left(Y_{i2}(0)=1\right)} = 1 - \frac{\Pr\left(Y_{i2}\left(1,Y_{i1}(1)\right)=1\right)}{\Pr\left(Y_{i2}\left(0,Y_{i1}(0)\right)=1\right)} = 1 - \frac{\Pr\left(Y_{i2}\left(1,Y_{i1}(1)\right)=1\right)}{\Pr\left(Y_{i2}\left(0,Y_{i1}(1)\right)=1\right)} \frac{\Pr\left(Y_{i2}\left(0,Y_{i1}(1)\right)=1\right)}{\Pr\left(Y_{i2}\left(0,Y_{i1}(0)\right)=1\right)} \qquad (2.5)$$

$$= 1 - (1-I_U)(1-C) = I_U + C - I_U C.$$

This decomposition is analogous to what are referred to as "natural direct and indirect effects" [16,17] in the mediation analysis.

3. Relationship between Conditional and Unconditional Infectiousness Effects

Here, we require the following two assumptions:

Assumption 1. Only person 1, not person 2, can become infected from outside the household.

Assumption 2. $Y_{i1}(1) \le Y_{i1}(0)$ for all i.

Assumption 1 implies that person 2 cannot be infected unless person 1 is infected. Assumption 2 implies $\Pr\left(Y_{i1}(1)=1, Y_{i1}(0)=0\right)=0$; *i.e.*, there is no household in which person 1 would be infected if vaccinated, but uninfected if unvaccinated.

Under these two assumptions, $\Pr\left(Y_{i2}\left(a_{i1}, Y_{i1}(1)\right)=1\right)$ can be expressed as follows [7]:

$$\Pr\left(Y_{i2}\left(a_{i1}, Y_{i1}(1)\right)=1\right) = \sum_{s=0}^{1}\sum_{t=0}^{1}\left\{\Pr\left(Y_{i2}\left(a_{i1}, Y_{i1}(1)\right)=1\middle|Y_{i1}(1)=s, Y_{i1}(0)=t\right) \times \Pr\left(Y_{i1}(1)=s, Y_{i1}(0)=t\right)\right\}$$

$$= \sum_{s=0}^{1}\sum_{t=0}^{1}\left\{\Pr\left(Y_{i2}\left(a_{i1}, s\right)=1\middle|Y_{i1}(1)=s, Y_{i1}(0)=t\right) \times \Pr\left(Y_{i1}(1)=s, Y_{i1}(0)=t\right)\right\}$$

$$= \Pr\left(Y_{i2}\left(a_{i1}, 1\right)=1\middle|Y_{i1}(1)=1, Y_{i1}(0)=1\right) \times \Pr\left(Y_{i1}(1)=1, Y_{i1}(0)=1\right)$$

$$= \Pr\left(Y_{i2}\left(a_{i1}, Y_{i1}(a_{i1})\right)=1\middle|Y_{i1}(1)=Y_{i1}(0)=1\right) \times \Pr\left(Y_{i1}(1)=1\right)$$

$$= \Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|Y_{i1}(1)=Y_{i1}(0)=1\right) \times \Pr\left(Y_{i1}=1\middle|A_{i1}=1\right),$$

where the third equality is because $\Pr\left(Y_{i2}\left(a_{i1},0\right)=1\middle|Y_{i1}(1)=0, Y_{i1}(0)=t\right)=0$ by Assumption 1 and $\Pr\left(Y_{i1}(1)=1, Y_{i1}(0)=0\right)=0$ by Assumption 2, the fourth is by Assumption 2, and the last is by random assignment to person 1 and the consistency assump-

tion.

Using this equation and Equations (2.2) and (2.3), it is readily confirmed that $I_U = I_C$ under Assumptions 1 and 2. Furthermore, substituting this equation into Equation (2.4) derives the following form for the contagion effect:

$$C = 1 - \frac{\left\{\Pr\left(Y_{i2}(0)=1\middle|Y_{i1}(1)=Y_{i1}(0)=1\right) \times \Pr\left(Y_{i1}=1\middle|A_{i1}=1\right)\right\}}{\Pr\left(Y_{i2}=1\middle|A_{i1}=0\right)},$$

where the denominator is by

$$\Pr\left(Y_{i2}\left(0, Y_{i1}(0)\right)=1\right)$$

$$= \Pr\left(Y_{i2}(0)=1\right) = \Pr\left(Y_{i2}=1\middle|A_{i1}=0\right).$$

Therefore, to assess the unconditional infectiousness and contagion effects, we can apply the statistical methods developed for the conditional infectiousness effect. Such methods are described in Sections 4 and 5. We note that this form for the contagion effect is not its definition under the principal stratification approach.

4. Estimation

Section 4.1 presents the inverse-probability-weighting (IPW) estimators for the infectiousness and contagion effects. Section 4.2 shows that the analysis for the IPW can be implemented easily.

4.1. Estimators

Under Assumption 2, $\Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|Y_{i1}(1)=Y_{i1}(0)=1\right)$ can be expressed as follows:

$$\Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|Y_{i1}\left(1\right)=Y_{i1}\left(0\right)=1\right)=\Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|Y_{i1}\left(1\right)=1\right)=\Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|Y_{i1}\left(1\right)=1, A_{i1}=1\right)$$
$$=\Pr\left(Y_{i2}\left(a_{i1}\right)=1\middle|A_{i1}=1, Y_{i1}=1\right).$$

Specifically,
$$\Pr\left(Y_{i2}\left(1\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)$$
$$=\Pr\left(Y_{i2}=1\middle|A_{i1}=1, Y_{i1}=1\right)$$
by the consistency assumption. Thus, under Assumptions

1 and 2, the respective infectiousness and contagion effects can be expressed as:

$$I_C = I_U = 1 - \frac{\Pr\left(Y_{i2}=1\middle|A_{i1}=1, Y_{i1}=1\right)}{\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)}, \quad (4.1)$$

$$C = 1 - \frac{\left\{\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)\times\Pr\left(Y_{i1}=1\middle|A_{i1}=1\right)\right\}}{\Pr\left(Y_{i2}=1\middle|A_{i1}=0\right)}. \quad (4.2)$$

To derive estimators of $\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)$, we require the following assumption:

Assumption 3. $Y_{i2}\left(a_{i1}\right)$ is independent of A_{i1}, conditional on Y_{i1} and X_i.

This assumption states that all baseline covariates that

affect the vaccine status of both persons 1 and 2 are measured, and implies that, in **Figure 2**, all paths from A_1 to Y_2 are blocked except for the direct path $A_1 \rightarrow Y_2$.

Under Assumption 3, $\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)$ can be expressed as follows [18]:

$$\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)=\sum_x\left\{\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1, X_i=x\right)\times\Pr\left(X_i=x\middle|A_{i1}=1, Y_{i1}=1\right)\right\}$$
$$=\sum_x\left\{\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=0, Y_{i1}=1, X_i=x\right)\times\Pr\left(X_i=x\middle|A_{i1}=1, Y_{i1}=1\right)\right\}$$
$$=\sum_x\left\{\Pr\left(Y_{i2}=1\middle|A_{i1}=0, Y_{i1}=1, X_i=x\right)\times\Pr\left(X_i=x\middle|A_{i1}=1, Y_{i1}=1\right)\right\}$$
$$=\frac{\sum_x p_{ix}\Pr\left(Y_{i2}=1, A_{i1}=0, Y_{i1}=1, X_i=x\right)\big/\left(1-p_{ix}\right)}{\Pr\left(A_{i1}=1, Y_{i1}=1\right)},$$

where $p_{ix}=Pr\left(A_{i1}=1\middle|Y_{i1}=1, X_i=x\right)$. Following the theory of Hirano *et al.* [19], once p_{ix} has been modeled and calculated, the last equation can be estimated by:

$$\frac{1}{N_{11}}\sum_{i=1}^N I\left(A_{i1}=0, Y_{i1}=1\right)\frac{p_{ix}}{1-p_{ix}}Y_{i2}, \quad (4.3)$$

where N_{11} is the number of households with $\left(A_{i1}, Y_{i1}\right)=\left(1, 1\right)$, and $I\left(.\right)$ denotes the indicator function with $I\left(A_{i1}=0, Y_{i1}=1\right)=1$ for households with $\left(A_{i1}, Y_{i1}\right)=\left(0, 1\right)$ and $I\left(A_{i1}=0, Y_{i1}=1\right)=0$ for the others. The value of p_{ix} is often predicted from a model for the regression of A_{i1} on X_i (*e.g.*, logistic regression model) in the subgroup of households with $Y_{i1}=1$.

We note that we can derive the other types of estimators; *i.e.*, the model-based standardization estimator [20], which uses a model for the regression of Y_{i2} on X_i rather than the regression of A_{i1} on X_i, and the doubly robust estimator [21], which uses both regression models. See Chiba and Taguri [18] for details.

4.2. Implementation

Equation (4.3) implies that we can implement the analysis by limiting the analysis set to households with $Y_{i1}=1$. Furthermore, except conditioning on $Y_{i1}=1$, Equation

(4.3) has the same form as those for the average causal effect of an exposure on the outcome with the exposed group as the target population in the setting of observational studies [20,22], where the exposure and outcome correspond to A_{i1} and Y_{i2}, respectively. Therefore, we can easily calculate the IPW estimate and the confidence interval (CI) using a marginal structural model (MSM) [23,24]. The regression parameters in this model are estimated by a weighted regression model with the form $Y_{i2}=\exp\left(\beta_0+\beta_1 A_{i1}\right)$, where the weights are $w_i=1$ for households with $A_{i1}=1$ and $w_i=p_{ix}/\left(1-p_{ix}\right)$ for households with $A_{i1}=0$, where the variance is estimated by the robust variance, which provides a conservative CI that is guaranteed to cover the true at least 95% of the time in large samples [24]. The SAS code is given elsewhere [7,18].

In this MSM, $\hat{\beta}_1$ corresponds to an estimator of

$$\log\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right)$$
$$-\log\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right),$$

and $\hat{\beta}_0$ corresponds to an estimator of
$$\log\Pr\left(Y_{i2}\left(0\right)=1\middle|A_{i1}=1, Y_{i1}=1\right).$$

Therefore, an estimator of the infectiousness effect is:

$$\hat{I}_C = \hat{I}_U = 1 - e^{\hat{\beta}_1}, \qquad (4.4)$$

and that of the contagion effect is:

$$\hat{C} = 1 - e^{\hat{\beta}_0} \frac{\hat{\Pr}(Y_{i1} = 1 | A_{i1} = 1)}{\hat{\Pr}(Y_{i2} = 1 | A_{i1} = 0)}. \qquad (4.5)$$

The CI of $\hat{I}_C = \hat{I}_U$ is evaluated by $1 - \exp\{\text{CI of } \hat{\beta}_1\}$, and similarly the CI of $\hat{C}$ is evaluated by

$$1 - \exp\{\text{CI of } \hat{\beta}_0 + \log \hat{\Pr}(Y_{i1} = 1 | A_{i1} = 1)$$
$$- \log \hat{\Pr}(Y_{i1} = 1 | A_{i1} = 0)\}.$$

In other words, the variances of these estimates are used to obtain the CIs, where the delta method is used to yield the variances of estimates on a log-scale.

5. Sensitivity Analysis

Here, we provide a simple sensitivity analysis method to assess how inference would change under violation of Assumption 3. The sensitivity analysis formula is provided in Section 5.1, and a plausible range of the sensitivity parameter is provided in Section 5.2.

5.1. Sensitivity Analysis Formulas

We set the sensitivity parameter as [18,25]:

$$\alpha = \frac{\Pr(Y_{i2}(0) = 1 | A_{i1} = 1, Y_{i1} = 1)}{\Pr(Y_{i2}(0) = 1 | A_{i1} = 0, Y_{i1} = 1)}. \qquad (5.1)$$

The sensitivity parameter α is the ratio between the outcome that would have been observed if person 1 was unvaccinated in comparing two different populations: the population in the numerator is that in which person 1 was vaccinated ($A_{i1} = 1$), and the population in the denominator is that in which person 1 was unvaccinated ($A_{i1} = 0$), where the infection status is equal in these two populations ($Y_{i1} = 1$). The interpretation of α is then simply the ratio of the expected outcomes under non-vaccination for these two populations.

Using Equation (5.1), Equations (4.1) and (4.2) can be expressed as follows:

$$I_C = I_U = 1 - \frac{1}{\alpha} \frac{\Pr(Y_{i2} = 1 | A_{i1} = 1, Y_{i1} = 1)}{\Pr(Y_{i2} = 1 | A_{i1} = 0, Y_{i1} = 1)}, \qquad (5.2)$$

$$C = 1 - \alpha \frac{\Pr(Y_{i1} = 1 | A_{i1} = 1)}{\Pr(Y_{i1} = 1 | A_{i1} = 0)}, \qquad (5.3)$$

respectively, where Equation (5.3) can be derived because, by Assumption 1,

$$\Pr(Y_{i2} = 1 | A_{i1} = 0)$$
$$= \sum_{y=0}^{1} \Pr(Y_{i2} = 1 | A_{i1} = 0, Y_{i1} = y) \Pr(Y_{i1} = y | A_{i1} = 0)$$
$$= \Pr(Y_{i2} = 1 | A_{i1} = 0, Y_{i1} = 1) \Pr(Y_{i1} = 1 | A_{i1} = 0).$$

Using Equations (5.2) and (5.3), a sensitivity analysis can be conducted as follows. The sensitivity parameter α is set by the investigator according to what is considered plausible. The parameter can be varied over a range of plausible values to examine how conclusions vary according to differences in parameter values. The results of the sensitivity analysis can be displayed graphically, where the horizontal axis represents the sensitivity parameter and the vertical axis represents the true infectiousness and contagion effects.

Generally, to obtain the CIs, the variances are calculated from Equations (5.2) and (5.3) with a fixed value of α. However, this calculation yields narrower CIs than those calculated from Equations (4.4) and (4.5). To avoid this problem, we use the variances calculated from Equations (4.4) and (4.5) to obtain the CIs of the infectiousness and contagion effects estimated from Equations (5.2) and (5.3).

5.2. Range of the Sensitivity Parameter

In some situations, it may be troublesome for investigators to determine a range of values of α to examine. Therefore, we present a range of values that α can take; *i.e.*, bounds for α.

Initially, we apply the large sample bounds [26-28] for

$$\Pr(Y_{i2}(0) = 1 | A_{i1} = 1, Y_{i1} = 1)$$
$$= \Pr(Y_{i2}(0) = 1 | Y_{i1}(1) = Y_{i1}(0) = 1).$$

under Assumption 2. The large sample bounds are calculated from the expected number if households with $(Y_{i1}(1), Y_{i1}(0)) = (1, 1)$ were assigned to the unvaccinated group ($A_{i1} = 0$), as follows.

Households with $(A_{i1}, Y_{i1}) = (0,1)$ are those with either $(Y_{i1}(1), Y_{i1}(0)) = (1,1)$ or $(Y_{i1}(1), Y_{i1}(0)) = (0,1)$. Therefore, under Assumption 2, of households with $(A_{i1}, Y_{i1}) = (0,1)$, the expected number of households with $(Y_{i1}(1), Y_{i1}(0)) = (1,1)$ assigned to the unvaccinated group ($A_{i1} = 0$), N^*, is calculated as:

$$N^* = \frac{N \Pr(A_{i1} = 0, Y_{i1} = 1) \times \Pr(Y_{i1}(1) = Y_{i1}(0) = 1)}{\Pr(Y_{i1}(1) = Y_{i1}(0) = 1) + \Pr(Y_{i1}(1) = 0, Y_{i1}(0) = 1)} = \frac{N \Pr(Y_{i1} = 1 | A_{i1} = 0) \Pr(A_{i1} = 0) \Pr(Y_{i1}(1) = 1)}{\Pr(Y_{i1}(0) = 1)}$$

$$= \frac{N \Pr(Y_{i1} = 1 | A_{i1} = 0) \Pr(A_{i1} = 0) \Pr(Y_{i1} = 1 | A_{i1} = 1)}{\Pr(Y_{i1} = 1 | A_{i1} = 0)} = N \Pr(A_{i1} = 0) \Pr(Y_{i1} = 1 | A_{i1} = 1).$$

Using this number, the lower bound of

$$\Pr\left(Y_{i2}(0)=1\big|A_{i1}=1,Y_{i1}=1\right)$$
$$=\Pr\left(Y_{i2}(0)=1\big|Y_{i1}(1)=Y_{i1}(0)=1\right)$$

is

$$\max\left\{\frac{N^{*}\times0}{N^{*}},\frac{N_{010}\times0+\left(N^{*}-N_{010}\right)\times1}{N^{*}}\right\},$$

and the upper bound is

$$\min\left\{\frac{N^{*}\times1}{N^{*}},\frac{\left(N^{*}-N_{011}\right)\times0+N_{011}\times1}{N^{*}}\right\},$$

where $N_{ayy^{*}}=N\Pr\left(A_{i1}=a,Y_{i1}=y,Y_{i2}=y^{*}\right)$ is the number of households with $\left(A_{i1},Y_{i1},Y_{i2}\right)=\left(a,y,y^{*}\right)$. Some algebra yields the following bounds for α:

$$\max\left\{0,1-\left(\frac{1}{P}-1\right)\left(\frac{1}{Q}-1\right)\right\}\le\alpha\le\min\left\{\frac{1}{P},\frac{1}{Q}\right\},$$

where $P=\Pr\left(Y_{i1}=1\big|A_{i1}=1\right)\big/\Pr\left(Y_{i1}=1\big|A_{i1}=0\right)$ and $Q=\Pr\left(Y_{i2}=1\big|A_{i1}=0,Y_{i1}=1\right)$. We note that the upper bound is the same as or smaller than that derived by Chiba and Taguri [7], which is $1/P$ under Assumptions 1 and 2.

Next, we consider the bounds for α under Assumption 2 and the following assumption [25,29]:

Assumption 4.

$$\Pr\left(Y_{i2}(0)=1\big|A_{i1}=1,Y_{i1}=1\right)$$
$$\ge\Pr\left(Y_{i2}(0)=1\big|A_{i1}=0,Y_{i1}=1\right).$$

This assumption is arguably reasonable insofar as the population in which person 1 was vaccinated is likely to be less healthy (or the infection is more virulent) than that in which person 1 was unvaccinated. The first population in which person 1 was infected despite receiving the vaccination would be less healthy than the second population in which person 1 was unvaccinated and infected. Thus, under the scenario in which both persons are unvaccinated, person 2 is more likely to be infected in the first than in the second population.

Under Assumptions 2 and 4, it is obvious that $\alpha\ge1$. While the large sample bounds generally yield a wide range, if Assumption 4 is plausible, it will yield a narrower width than the large sample bounds only.

6. Illustration

To illustrate the methods presented in Sections 4 and 5, we employ data from a hypothetical vaccine trial used elsewhere [4]. Consider now a vaccine trial setting in which 1-year-old children at a day-care center are randomized to receive pneumococcal conjugate vaccine against a given pneumococcus serotype. The colonization

status with respect to the given serotype of the 1-year-old child and the mother is also monitored. Because pneumococcus is highly prevalent in young children who attend day care, the mother is more likely to acquire the pneumococcus from the child than through other transmission routes. Here, we assume that the second person (the mother) can be infected only from the first (the 1-year-old child). The hypothetical data are given in **Table 1**.

Under Assumptions 1-3, using a MSM, Equations (4.4) and (4.5) yielded an IPW estimate of the infectiousness effect of 0.44 (95% CI: 0.33, 0.53) and an IPW estimate of the contagion effect of 0.54 (95% CI: 0.46, 0.61). The overall indirect effect was $1-(79/1000)/(305/1000)=0.74$ (95% CI: 0.67, 0.79), which equals the decomposition of Equation (2.5): $0.44+0.54-0.44\times0.54=0.74$.

To assess how inference would change under violation of Assumption 3, we implemented the sensitivity analysis presented in Section 5. Before the implementation, we determined the bounds for α to determine the range to be examined. The large sample bounds yielded bounds for α of $0.04\le\alpha\le1.80$. By adding Assumption 4, this range was narrowed to $1.00\le\alpha\le1.80$. For this range of α, we implemented the sensitivity analysis using Equations (5.2) and (5.3). The results are shown in **Figure 3** for the infectiousness effect and **Figure 4** for the contagion effect.

For this range of α, the respective lower and upper limits of the infectiousness effect were 0.43 (95% CI: 0.31, 0.53) and 0.68 (95% CI: 0.62, 0.74), and those of the contagion effect were 0.18 (95% CI: 0.04, 0.30) and 0.55 (95% CI: 0.47, 0.61). The IPW estimates under Assumption 3 correspond to the estimates for $\alpha=1.01$ in **Figures 3** and **4**.

Table 1. Numbers infected (Y_{i1}, Y_{i2}) from a hypothetical randomized trial of pneumococcal conjugate vaccine with 2000 households[a].

	$Y_{i1}=0$ $Y_{i2}=0$	$Y_{i1}=1$ $Y_{i2}=0$	$Y_{i1}=1$ $Y_{i2}=1$	Total
Low SES				
$A_{i1}=0,A_{i2}=0$	200	120	180	500
$A_{i1}=1,A_{i2}=0$	350	96	54	500
High SES				
$A_{i1}=0,A_{i2}=0$	250	125	125	500
$A_{i1}=1,A_{i2}=0$	400	75	25	500

[a]Person 1 (the 1-year-old child) was randomized 1:1 to vaccine or control, and person 2 (the mother) was not vaccinated (vaccination status [A_{i1}, A_{i2}]), and half the households have either low or high socioeconomic status (SES).
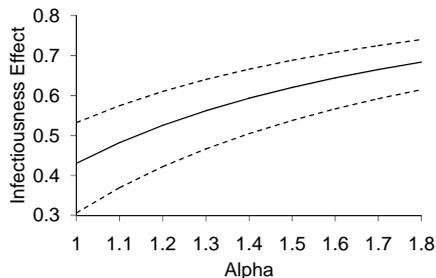
Figure 3. Sensitivity analysis of the infectiousness effect; the solid line indicates the infectiousness effect and broken lines indicate 95% confidence intervals.
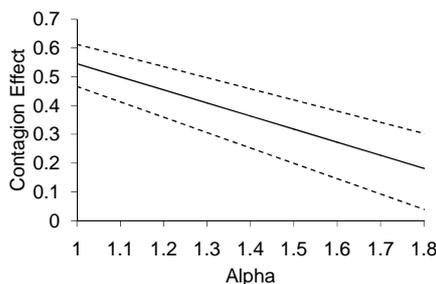

Figure 4. Sensitivity analysis of the contagion effect; the solid line indicates the contagion effect and broken lines indicate 95% confidence intervals.

7. Discussion

In this paper, by applying a simple relationship between conditional and unconditional infectiousness effects, we have presented simple statistical methods to assess the infectiousness and contagion effects. Although the methods for the infectiousness effects are a review of the past literature, we have summarized them in the present paper. We have further applied them to infer the contagion effect.

The methods presented here are limited by the assumptions. While Assumption 1 can be partially verified from the observed data by examining whether $\Pr\left(Y_{i2}=1 \mid Y_{i1}=0\right)=0$, Assumption 2 cannot be verified from the observed data. Although we believe that Assumption 2 will be plausible in many settings, whether the assumption holds will depend on the nature of the vaccine under study. Therefore, the assumption will not be applicable to all vaccines [18]. In such situations, unfortunately, the methods presented here cannot be applied. Nevertheless, for unconditional infectiousness and contagion effects, we can still apply the methods developed in the context of the mediation analysis approach [30-33], and for the conditional infectiousness effect, we can also apply the methods developed in the context of the principal stratification approach [25-28]. However, these methods have a weakness in that they are more complex than those in this paper.

The framework used also made an assumption of par-

tial interference. This might be plausible if the various households are sufficiently geographically separated or do not interact. In certain settings, this assumption might be plausible. Nevertheless, future work will attempt to generalize the methods given here to allow for violations of this assumption.

8. Acknowledgements

REFERENCES

[1] M. E. Halloran and M. G. Hudgens, "Causal Inference for Vaccine Effects on Infectiousness," *International Journal of Biostatistics*, Vol. 8, No. 2, 2012, Article 6. doi:10.2202/1557-4679.1354

[2] P. R. Rosenbaum, "Interference between Units in Randomized Experiments," *Journal of the American Statistical Association*, Vol. 102, No. 477, 2007, pp. 191-200. doi:10.1198/016214506000001112

[3] M. G. Hudgens and M. E. Halloran, "Towards Causal Inference with Interference," *Journal of the American Statistical Association*, Vol. 103, No. 482, 2008, pp. 832-842. doi:10.1198/016214508000000292

[4] T. J. VanderWeele, E. J. Tchetgen Tchetgen and M. E. Halloran, "Components of the Indirect Effect in Vaccine Trials: Identification of Contagion and Infectiousness Effects," *Epidemiology*, Vol. 23, No. 5, 2012, pp. 751-761. doi:10.1097/EDE.0b013e31825fb7a0

[5] S. Datta, M. E. Halloran and I. M. Longini, "Efficiency of Estimating Vaccine Efficacy for Susceptibility and Infectiousness: Randomization by Individual Versus Household," *Biometrics*, Vol. 55, No. 3, 1999, pp. 792-798. doi:10.1111/j.0006-341X.1999.00792.x

[6] M. G. Hudgens and M. E. Halloran, "Causal Vaccine Effects on Binary Post-Infection Outcomes," *Journal of the American Statistical Association*, Vol. 101, No. 473, 2006, pp. 51-64. doi:10.1198/016214505000000970

[7] Y. Chiba and M. Taguri, "Conditional and Unconditional Infectiousness Effects in Vaccine Trials," *Epidemiology*, Vol. 24, No. 2, 2013, pp. 336-337. doi:10.1097/EDE.0b013e31828261f5

[8] M. M. Glymour and S. Greenland, "Causal Diagrams," In: K. J. Rothman, S. Greenland and T. L. Lash, Eds., *Modern Epidemiology*, 3rd Edition, Lippincott Williams and Wilkins, Philadelphia, 2008, pp. 183-209.

[9] J. Pearl, "Causality: Models, Reasoning, and Inference," 2nd Edition, Cambridge University Press, Cambridge, 2009.

[10] M. E. Sobel, "What Do Randomized Studies of Housing Mobility Demonstrate? Causal Inference in the Face of Interference," *Journal of the American Statistical Association*, Vol. 101, No. 476, 2006, pp. 1398-1407.

doi:10.1198/016214506000000636

[11] T. J. VanderWeele, "Concerning the Consistency Assumption in Causal Inference," *Epidemiology*, Vol. 20, No. 6, 2009, pp. 880-883. doi:10.1097/EDE.0b013e3181bd5638

[12] M. E. Halloran, M. P. Préziosi and H. Chu, "Estimating Vaccine Efficacy from Secondary Attack Rates," *Journal of the American Statistical Association*, Vol. 98, No. 461, 2003, pp. 38-46. doi:10.1198/016214503388619076

[13] S. Greenland, J. Pearl and J. M. Robins, "Causal Diagrams for Epidemiologic Research," *Epidemiology*, Vol. 10, No. 1, 1999, pp. 37-48.

[14] M. G. Hudgens and M. E. Halloran, "Causal Vaccine Effects on Binary Post-Infection Outcomes," *Journal of the American Statistical Association*, Vol. 101, No. 473, 2006, pp. 51-64. doi:10.1198/016214505000000970

[15] C. E. Frangakis and D. B. Rubin, "Principal Stratification in Causal Inference," *Biometrics*, Vol. 58, No. 1, 2002, pp. 21-29. doi:10.1111/j.0006-341X.2002.00021.x

[16] J. M. Robins and S. Greenland, "Identifiability and Exchangeability for Direct and Indirect Effects," *Epidemiology*, Vol. 3, No. 2, 1992, pp. 143-155.

[17] J. Pearl, "Direct and Indirect Effects," *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*, San Francisco, 2-5 August 2001, pp. 411-420.

[18] Y. Chiba and M. Taguri, "Assessing the Causal Infectiousness Effect in Vaccine Trials," In: iConcept Press, Ed., *Vaccines—Benefits and Risks*, iConcept Press, Hong Kong, 2013, in press.

[19] K. Hirano, G. W. Imbens and G. Ridder, "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, Vol. 71, No. 4, 2003, pp. 1161-1189.

[20] S. Greenland, "Estimating Standardized Parameters from Generalized Linear Models," *Statistics in Medicine*, Vol. 10, No. 7, 1991, pp. 1069-1074. doi:10.1002/sim.4780100707

[21] M. Taguri, Y. Matsuyama, Y. Ohashi, A. Harada and H. Ueshima, "Doubly Robust Estimation of the Generalized Impact Fraction," *Biostatistics*, Vol. 13, No. 3, 2012, pp. 455-467. doi:10.1093/biostatistics/kxr038

[22] T. Sato and Y. Matsuyama, "Marginal Structural Models as a Tool for Standardization," *Epidemiology*, Vol. 14, No. 6, 2003, pp. 680-686. doi:10.1097/01.EDE.0000081989.82616.7d

[23] J. M. Robins, "Association, Causation, and Marginal

Structural Models," *Synthese*, Vol. 121, No. 1-2, 1999, pp. 151-179. doi:10.1023/A:1005285815569

[24] J. M. Robins, M. A. Hernán and B. A. Brumback, "Marginal Structural Models and Causal Inference in Epidemiology," *Epidemiology*, Vol. 11, No. 5, 2000, pp. 550-560.

[25] Y. Chiba and T. J. VanderWeele, "A Simple Method for Principal Strata Effects When the Outcome Has Been Truncated Due to Death," *American Journal of Epidemiology*, Vol. 173, No. 7, 2011, pp. 745-751. doi:10.1093/aje/kwq418

[26] J. L. Zhang and D. B. Rubin, "Estimation of Causal Effects via Principal Stratification When Some Outcomes Are Truncated by Death," *Journal of Educational and Behavioral Statistics*, Vol. 28, No. 4, 2003, pp. 353-368. doi:10.3102/10769986028004353

[27] J. L. Zhang, D. B. Rubin and F. Mealli, "Evaluating the Effects of Job Training Programs on Wages through Principal Stratification," In: D. Millimet, J. Smith and E. Vytlacil, Eds., *Advances in Econometrics*: *Modeling and Evaluating Treatment Effects in Econometrics*, Elsevier Science, New York, 2008, pp. 117-145.

[28] Y. Chiba, "The Large Sample Bounds on the Principal Strata Effect with Application to a Prostate Cancer Prevention Trial," *International Journal of Biostatistics*, Vol. 8, No. 1, 2012, Article 12. doi:10.1515/1557-4679.1365

[29] T. J. VanderWeele and E. J. Tchetgen Tchetgen, "Bounding the Infectiousness Effect in Vaccine Trials," *Epidemiology*, Vol. 22, No. 5, 2011, pp. 686-693. doi:10.1097/EDE.0b013e31822708d5

[30] Sjölander, "Bounds on Natural Direct Effects in the Presence of Confounded Intermediate Variables," *Statistics in Medicine*, Vol. 28, No. 4, 2009, pp. 558-571. doi:10.1002/sim.3493

[31] T. Lange, S. Vansteelandt and M. Bekaert, "A Simple Unified Approach for Estimating Natural Direct and Indirect Effects," *American Journal of Epidemiology*, Vol. 176, No. 3, 2012, pp. 190-195. doi:10.1093/aje/kwr525

[32] L. Valeri and T. J. VanderWeele, "Mediation Analysis Allowing for Exposure-Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation with SAS and SPSS Macros," *Psychological Methods*, Vol. 18, No. 2, 2013, pp. 137-150. doi:10.1037/a0031034

[33] Y. Chiba and M. Taguri, "Alternative Monotonicity Assumptions for Improving Bounds on Natural Direct Effects," *International Journal of Biostatistics*, in press. doi:10.1515/ijb-2012-0022