Risk ratios for contagious outcomes

Olga Morozova[1], Ted Cohen[1] and Forrest W. Crawford[2,3,4]

[1]Department of Epidemiology of Microbial Diseases, and [2]Department of Biostatistics, Yale School of Public Health, 60 College Street, New Haven, CT 06510, USA
[3]Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect St, New Haven, CT 06511, USA
[4]Yale School of Management, 165 Whitney Ave, New Haven, CT 06511, USA

OM, 0000-0001-7606-8413; TC, 0000-0002-8091-7198; FWC, 0000-0002-0046-0547

Epidemiologists commonly use the risk ratio to summarize the relationship between a binary covariate and outcome, even when outcomes may be dependent. Investigations of transmissible diseases in clusters—households, villages or small groups—often report risk ratios. Epidemiologists have warned that risk ratios may be misleading when outcomes are contagious, but the nature of this error is poorly understood. In this study, we assess the meaning of the risk ratio when outcomes are contagious. We provide a mathematical definition of infectious disease transmission within clusters, based on the canonical stochastic susceptible–infective model. From this characterization, we define the individual-level ratio of instantaneous infection risks as the inferential target, and evaluate the properties of the risk ratio as an approximation of this quantity. We exhibit analytically and by simulation the circumstances under which the risk ratio implies an effect whose direction is opposite that of the true effect of the covariate. In particular, the risk ratio can be greater than one even when the covariate reduces both individual-level susceptibility to infection, and transmissibility once infected. We explain these findings in the epidemiologic language of confounding and Simpson's paradox, underscoring the pitfalls of failing to account for transmission when outcomes are contagious.

1. Introduction

Risk ratios are often recommended for summarizing the relationship between a covariate and an outcome in epidemiology [1–10]. They are simple and easy to compute [11–13], and when outcomes may exhibit correlation within clusters, 'robust' standard errors are available [10,14]. Many researchers report risk ratios in studies of infectious disease outcomes within clusters or communities of interacting individuals, including studies of H1N1 influenza [15,16], Ebola [17,18], leprosy [19] and varicella [20]. When the outcome is an infection indicator, the covariate of interest may correspond to a preventive treatment like vaccination, a risk factor like immune status or the presence of a co-morbid condition, or a demographic characteristic like age, sex or socio-economic status. Contagion within clusters may induce dependence in cluster members' infection outcomes [21], but the effect of a covariate under contagion may be complicated: different values might alter an individual's susceptibility to infection, or their infectiousness once infected, or both [22–24].

Infectious disease epidemiologists have repeatedly warned that when outcomes are contagious, simplistic summaries of risk may be misleading [24–35]. Researchers have argued that bias can arise when analytical methods do not separate the effects of a covariate on susceptibility to infection from infectiousness once infected [24,29,36], or when individual-level variation in exposure to infection is ignored [26,27,34,36–39]. Epidemiologists have suggested that risk ratios may not always give a reasonable summary of individual-level covariate effects. For example, risk ratios are often assumed to be time-invariant, but they can change over time in both observational [40] and experimental [41] studies. Likewise the risk ratio, as a measure of the effect of a vaccine, depends on both the time at which outcomes are observed and the population-level vaccine coverage [42]. However, epidemiologists have not

explained how these issues are related, nor demonstrated formally why the risk ratio may be an unsatisfactory measure of association under contagion for empirical research on infectious diseases. Prior work does not provide guidance about the disease dynamics and study design features that are most likely to give rise to profoundly biased estimates.

In this paper, we investigate the properties of the risk ratio when outcomes are contagious within clusters, and explain how these properties depend on the covariate of interest and the epidemiologic features of disease transmission. We first introduce a canonical definition of infectious disease contagion, based on the widely used susceptible–infective epidemic model [43,44]. This structural description of disease transmission formalizes the epidemiologic intuition that a susceptible individual's risk of infection at a given time depends both on their own traits, and those of their infectious contacts [36,39,45]. We define the inferential target as the ratio of instantaneous individual-level risks (hazards) of infection under a one-unit change in the value of a covariate [28,34], and show that the risk ratio can be a profoundly biased approximation of this quantity. For the simplest setting of clusters of size two, we describe conditions under which the risk ratio exhibits the most egregious form of error, bias 'across the null', in which the direction of the estimated effect is in error. Further analytic results and simulations provide insight into the behaviour of the risk ratio under contagion in clusters of larger size and in randomized trials. These results provide the first formal description of the pathological properties of the risk ratio under common assumptions about contagion. Finally, we explain these findings in the epidemiologic language of bias induced by confounding.

2. Setting

Consider a collection of clusters (e.g. households, workplaces, villages), with n_i subjects in cluster i. Let $Y_{ij}(t)$ be the binary indicator of infection for subject j in cluster i on or before time $t \geq 0$. Let T_i be the time at which outcomes in cluster i are observed and recorded by researchers. We consider a single time-invariant binary covariate x_{ij} for subject j in cluster i. The risk ratio is defined as

$$\text{RR} = \frac{\mathbb{E}[Y_{ij}(T_i) \mid x_{ij} = 1]}{\mathbb{E}[Y_{ij}(T_i) \mid x_{ij} = 0]}. \tag{2.1}$$

The risk ratio is implicitly a function of the observation time T_i for each cluster i [46].

We describe a stochastic transmission model based on the canonical susceptible–infective model of infectious disease contagion within clusters [36,39,43–45], then characterize the hazard ratio corresponding to a one-unit change in a covariate associated with susceptibility to disease. The susceptible–infective model captures the intuition that the risk to a susceptible individual at a given time is given heuristically by

risk of infection = (susceptibility) × (force of infection), (2.2)

where 'susceptibility' is a function of the subject's own characteristics, and 'force of infection' summarizes the risk transmitted by that subject's infectious contacts, including from outside of the cluster.

To formalize this risk, let t_{ij} be the minimum of the infection time of subject j in cluster i and the observation time T_i, so that $Y_{ij}(t) = 0$ for $t \leq t_{ij}$, and $Y_{ij}(t) = 1$ for $t_{ij} < t \leq T_i$. A subject j in cluster i is called susceptible at time t if $Y_{ij}(t) = 0$, and infected if $Y_{ij}(t) = 1$. Consider the possible sources of transmission to a susceptible subject j in cluster i. First, j may be infected by exposure to an exogenous source of infection (sometimes called the community force of infection if clusters are households). Let τ_{ij}^e be the waiting time for j to be infected from this exogenous source, and let $\lambda_{ij}^e(t)$ be the hazard of this event at time t. Second, suppose another subject k in cluster i becomes infected at a time $t_{ik} \in [0, T_i)$, which is defined similarly to t_{ij} as the minimum of the infection time of subject k and cluster i observation time T_i. Suppose subject j is not infected at time t_{ik}, $Y_{ij}(t_{ik}) = 0$. Let τ_{ij}^k be the waiting time (measured since t_{ik}) for k to transmit the infection to j, and let $\lambda_{ij}^k(t)$ be the hazard of this event at time $t > t_{ik}$. For each cluster i and susceptible subject j, the total hazard experienced by a susceptible individual j is the sum of these hazards,

$$\lambda_{ij}(t) = \lambda_{ij}^e(t) + \sum_{k=1}^{n_i} \lambda_{ij}^k(t) Y_{ik}(t). \tag{2.3}$$

The additive form of (2.3) arises because j experiences *competing risks* of infection: from the exogenous source, and from each of their infectious contacts. Under this simple generative process, subjects may not be re-infected.

We assume for simplicity that the hazards $\lambda_{ij}^e(t)$ and $\lambda_{ij}^k(t)$ are Cox-type models: each decomposes into the product of a possibly time-varying force of infection and a function of covariates. Let $\lambda_{ij}^e(t) = \alpha_i(t) \exp[x_{ij}\beta]$, where $\alpha_i(t)$ is the possibly time-varying exogenous force of infection to cluster i, and β is a susceptibility parameter corresponding to the binary covariate x. Likewise, when $t > t_{ik}$, let $\lambda_{ij}^k(t) = \omega_{ikj}(t - t_{ik}) \exp[x_{ij}\beta + x_{ik}\gamma]$ where $\omega_{ikj}(t - t_{ik})$ is the possibly time-varying force of infection from subject k to subject j in cluster i, and γ is an infectiousness parameter corresponding to the binary covariate x. Then the total infection hazard to susceptible subject j in cluster i at time t becomes

$$\lambda_{ij}(t) = e^{x_{ij}\beta}\left(\alpha_i(t) + \sum_{k=1}^{n_i} Y_{ik}(t)\omega_{ikj}(t - t_{ik}) e^{x_{ik}\gamma}\right). \tag{2.4}$$

The multiplicative relationship between susceptibility $\exp[x_{ij}\beta]$ and the total force of infection in (2.4) mirrors the heuristic description of infection risk given by (2.2). Figure 1 shows a schematic depiction of the transmission model in a cluster of size three. The risk of infection experienced by a susceptible subject j increases over time as individuals around j become infected. Since a covariate x may, in general, alter infectiousness, the magnitude of risk increase to j with every subsequent contact, who become infected, depends on that individual's covariate.

When $x_{ij} = x$ is constant across individuals, $\alpha_i(t) = 0$, and $\omega_{ikj}(t - t_{ik}) = \omega$, the process becomes the standard continuous-time Markov susceptible–infective model within clusters. The formulation of the hazard of infection in (2.4) mirrors a transmission model proposed for semi-parametric relative risk regression [36]. The model captures temporal changes in post-infection transmission via the functional form of $\omega_{ikj}(t - t_{ik})$, which can accommodate latency or other changes in infectiousness over time.
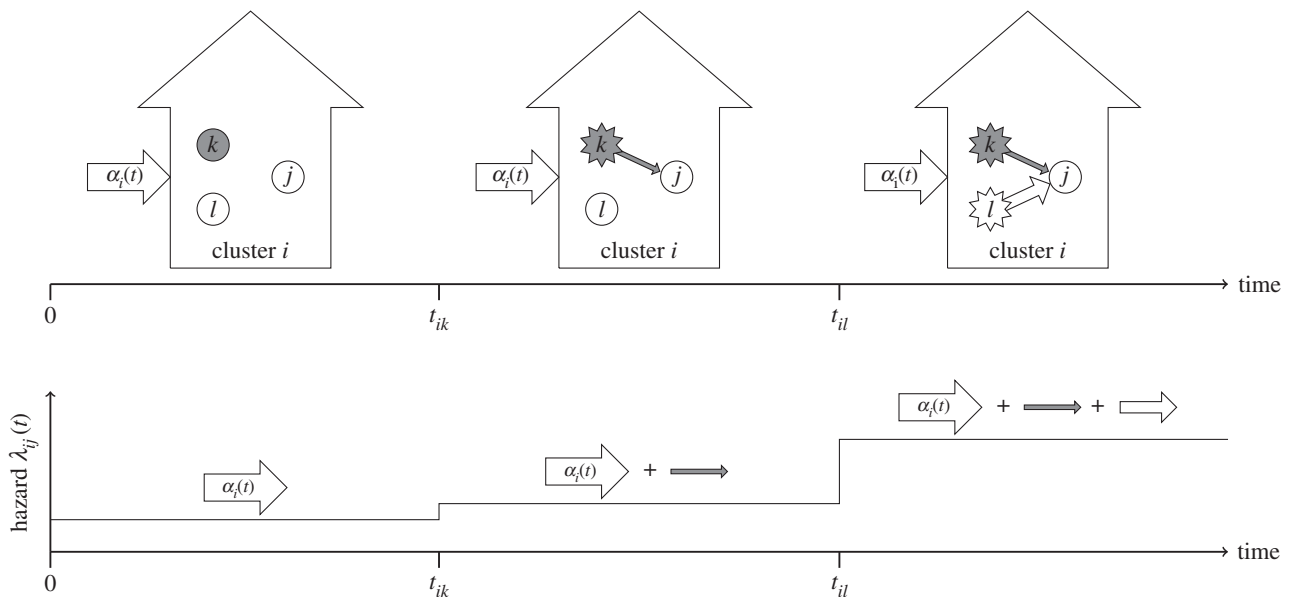
Figure 1. Schematic of hazards in the transmission model for a cluster of size three. Grey colour indicates $x = 1$. Before the first infection, subject j experiences only an exogenous (community) force of infection $\alpha_i(t)$, because neither k nor l is infected. After k is infected, the hazard to j increases in proportion to the infectiousness of k, which is a function of $x_{ik} = 1$. Likewise, after l is infected, the hazard to j increases in proportion to the infectiousness of l, with $x_{il} = 0$. Below, the sum of hazards experienced by subject j is shown over time.

Table 1. Summary of parameters in the transmission model.

notation	definition
n_i	size of cluster i
T_i	observation time for cluster i
t_{ij}	minimum of infection time of subject j in cluster i and cluster i observation time T_i
x_{ij}	covariate of interest, time-invariant
$Y_{ij}(t)$	binary indicator of infection by time t
β	susceptibility parameter for covariate x
γ	infectiousness parameter for covariate x
$\alpha_i(t)$	exogenous force of infection, a function of time
$\omega_{ikj}(t)$	force of infection from infectious k to susceptible j, a function of time since infection of k

The hazard ratio (HR) is the ratio of instantaneous risks given by (2.4) under different values of the covariate x, holding individual-level force of infection constant:

$$\text{HR} = \frac{\lambda_{ij}(t \mid x_{ij} = 1)}{\lambda_{ij}(t \mid x_{ij} = 0)} = e^{\beta}. \tag{2.5}$$

The hazard ratio summarizes the individual-level association between the covariate x and susceptibility to infection at time t [28,34,46].

We emphasize that we do not treat the transmission model characterized by (2.4) as an inferential model. We have not specified the possibly time-varying hazards $\alpha_i(t)$ and $\omega_{ikj}(t)$, nor showed that any feature of the process is identified by a particular observation scenario. Instead, (2.4) characterizes the transmission dynamics of infection by which the observable data are assumed to be generated. Table 1 summarizes the parameters that define this process.

It seems reasonable to expect the risk ratio given by (2.1) for the binary variable x, as a *marginal* or *population-level* measure of association, to be meaningful for assessment of the ratio of *conditional* risks (2.5) experienced by an individual. Since the hazard ratio (2.5) evaluated at a time t is time-invariant, we might expect the risk ratio, as a cross-sectional measure of association at time t, to provide similar evidence about the relationship between x and the outcome. In particular, researchers may wish to avoid a particularly egregious form of bias, in which the direction of the estimated effect is opposite that of the true effect. For example, if $x = 1$ is associated with reduced susceptibility to infection, we would like to see that RR < 1. To make this notion more formal, we define a general property that we would like the risk ratio to satisfy.

Definition 2.1. (Direction-unbiasedness of risk ratio.) If HR < 1, then RR < 1, if HR = 1, then RR = 1, and if HR > 1, then RR > 1.

We say that for a particular study design and values of parameters in (2.4), the risk ratio is *direction-unbiased* if definition 2.1 holds. When definition 2.1 does not hold, we say that the risk ratio exhibits 'bias across the null' because its direction is opposite that of the true effect e^{β}. Definition 2.1 is a relatively weak requirement: it does not imply zero bias, nor monotonicity in the risk ratio as a function of the hazard ratio, nor any particular functional relationship between the two.

3. Results

3.1. Clusters of size two

We first consider a simple parametric version of (2.4) with two-person clusters and balanced covariate values for which a variety of precise analytic results can be derived. This setting is based on a two-person infectious disease contagion model introduced previously [47–49], and serves to illustrate the potential for the risk ratio to give a misleading summary of
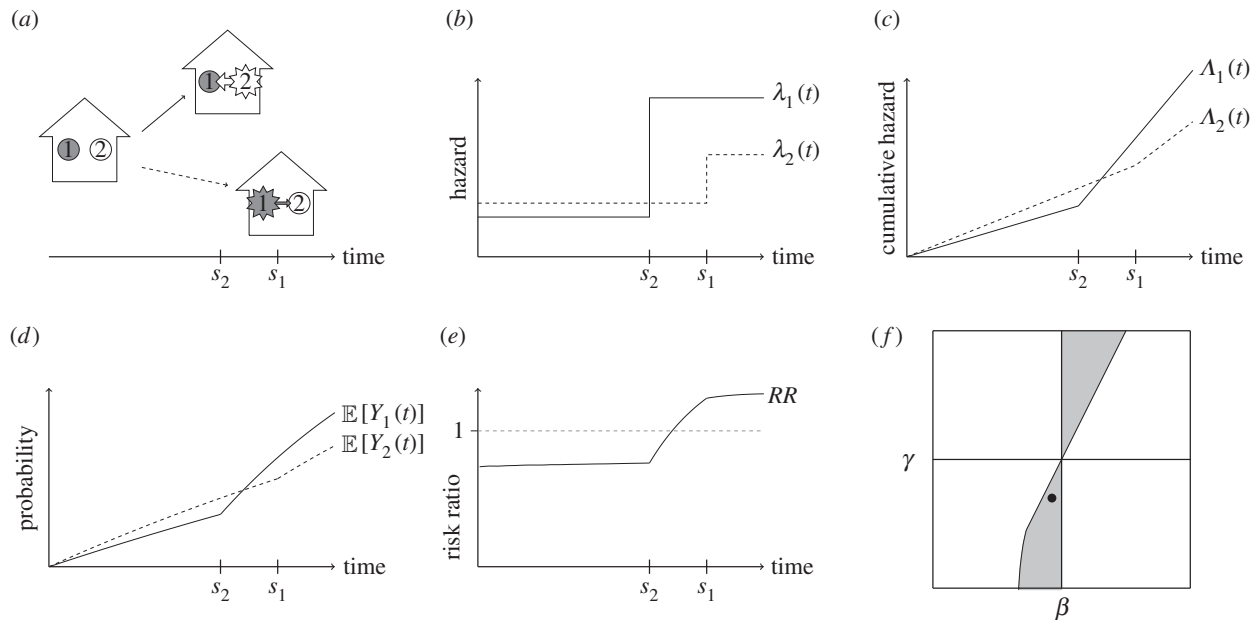
Figure 2. Illustration of how the risk ratio can give misleading results under contagion in a cluster of size two. Suppose the conditions of result 3.4 hold with $\beta < 0$ and $\gamma < 0$ (such that the presence of the covariate, designated as value $= 1$, is associated with reduced susceptibility and transmissibility). Subject 1 (grey) has $x_1 = 1$, subject 2 (white) has $x_2 = 0$. (a) Either subject 2 becomes infected first (at time s_2), or subject 1 is infected first (at time s_1); depending on which is infected first, the other experiences a change in their hazard of infection. (b) If 1 is infected first, then 2 experiences only a small increase in hazard, because $\gamma < 0$. Alternatively, if 2 is infected first, then 1 experiences a large increase in hazard because $x_2 = 0$. (c) The relationship between the cumulative hazards in these scenarios, and hence the relationship between the expected infection outcomes (d), is eventually reversed at some time after s_2. Therefore, the risk ratio (e) eventually rises above one. Panel (f) shows the region of (β, γ) parameter space in which bias across the null may occur, where β and γ are plotted on the same scale. A black dot indicates the values of β and γ used in this illustration.

association under contagion. Clusters of size two appear in empirical study designs, including HIV transmission in couples [50,51], and mother-to-child transmission of *Staphylococcus aureus* [52,53]. Consider the transmission model characterized by (2.4), where each cluster i consists of exactly two subjects: $n_i = 2$. Assume also that the covariate is balanced within the cluster, subject 1 has $x_{i1} = 1$ and subject 2 has $x_{i2} = 0$; all subjects are uninfected at baseline, $Y_{ij}(0) = 0$; and follow-up time is constant, $T_i = T$ for all i. Additionally, assume there is a constant exogenous force of infection $\alpha_i(t) = \alpha$, and constant within-cluster contagion $\omega_{ikj}(t - t_{ik}) = \omega$ per susceptible j and infected k. The hazards of infection experienced by subjects 1 and 2 in cluster i become $\lambda_{i1}(t) = \exp(\beta)(\alpha + \omega Y_{i2}(t))$ and $\lambda_{i2}(t) = \alpha + \omega \exp(\gamma) Y_{i1}(t)$, respectively. The following results establish the properties of the risk ratio as an approximation to the hazard ratio in several relevant special cases. Unless otherwise stated, we assume $\alpha > 0$ and $\omega > 0$. Derivations and proofs of all results are given in the electronic supplementary material.

Result 3.1. (No within-cluster contagion.) Suppose $\omega = 0$. For any $T > 0$, the risk ratio is direction-unbiased.

Thus when the outcome is not transmissible within clusters, the risk ratio is direction-unbiased.

Define the 'null' hypothesis under the transmission model as $\beta = 0$, so that all subjects are equally susceptible to infection.

Result 3.2. (Under the null.) Suppose $\beta = 0$ and $T > 0$. Then if $\gamma < 0$, RR > 1; if $\gamma > 0$, RR < 1; and if $\gamma = 0$, RR $= 1$.

This result means that when the covariate does not affect susceptibility to infection, the direction of the risk ratio is entirely determined by the infectiousness coefficient γ. In particular,

when the covariate reduces infectiousness ($\gamma < 0$)—as many treatments do—the risk ratio can be greater than one.

Result 3.3. (Homogeneous infectiousness.) Suppose $\gamma = 0$. For any $T > 0$, the risk ratio is direction-unbiased.

Therefore, if the covariate does not alter infectiousness, direction-unbiasedness holds.

Result 3.4. (Bias across the null.) Suppose either $\beta < 0$ and $e^{\gamma} < \min\{e^{2\beta}, e^{\beta} + (\alpha/\omega)(e^{\beta} - 1)\}$, or $\beta > 0$ and $e^{\gamma} > \max\{e^{2\beta}, e^{\beta} + (\alpha/\omega)(e^{\beta} - 1)\}$. Then there exists $t^* > 0$ such that for all $T > t^*$, the risk ratio is not direction-unbiased.

This result states that the risk ratio can be biased across the null when β is non-zero. Figure 2 illustrates result 3.4.

Direction-unbiasedness under definition 2.1 does not imply zero bias. Figure 3 shows log[RR] as a function of β and γ for several values of ω/α. Under the transmission model (2.4), any covariate can be represented by a point in this two-dimensional (β, γ) space, corresponding to its effects on susceptibility and infectiousness. Other study design parameters govern the behaviour of the risk ratio as a function of β and γ. The ratio ω/α summarizes the relative contribution of within-cluster transmission compared to transmission from the community. To make results comparable in every sub-figure, the observation time T is selected so that cumulative incidence at time T when $\beta = 0$ and $\gamma = 0$ is held constant at approximately 0.15. Figure 3a shows the magnitude of bias; as the ratio ω/α increases, the absolute value of the bias becomes larger even in regions where direction-unbiasedness holds. The electronic supplementary material provides a derivation of an exact expression for log[RR] and similar plots for a wider range of parameters α and ω. As an approximation to
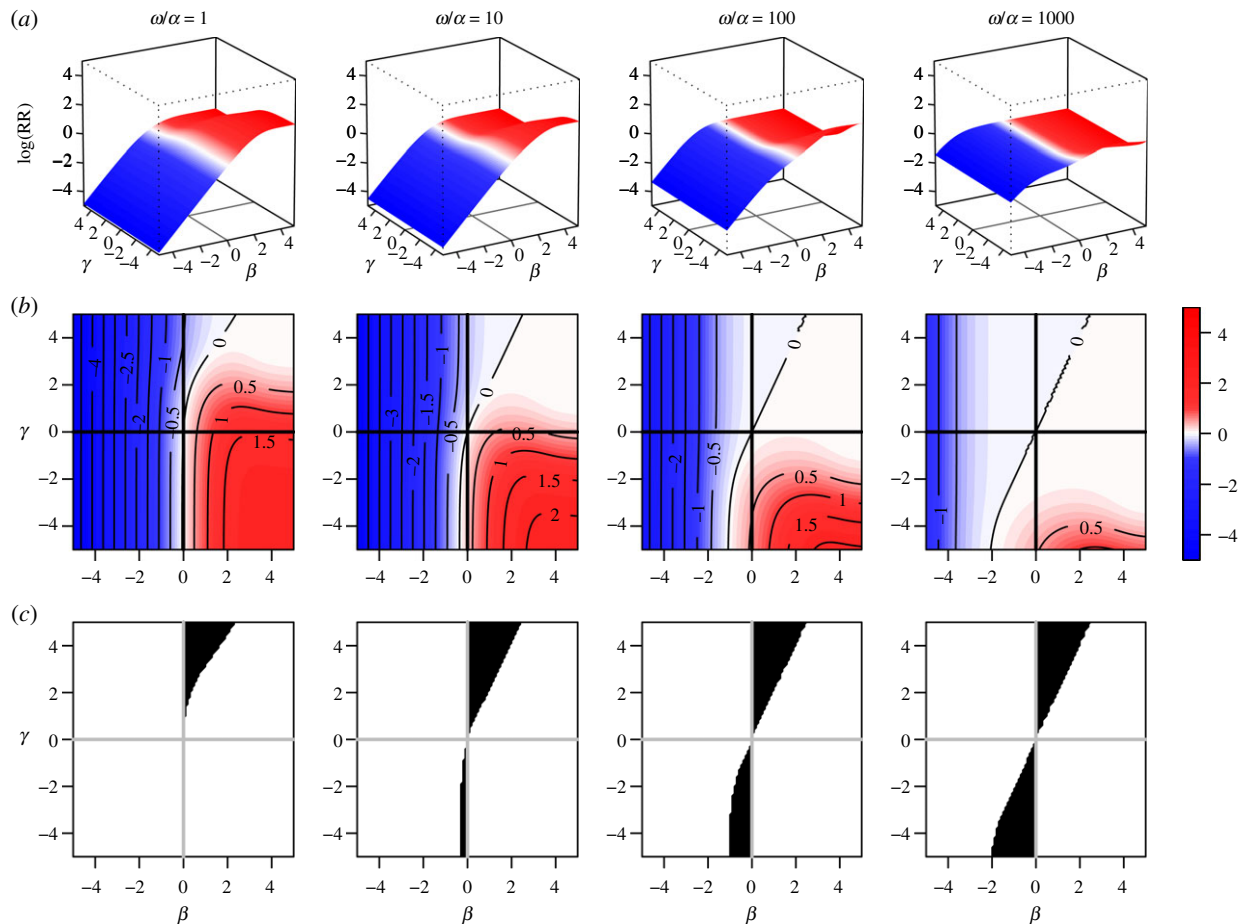
Figure 3. Risk ratio (log) as a function of true covariate effects β and γ. Row (*a*) shows the log risk ratio surface on (β, γ) space, and (*b*) shows the corresponding heat maps and contours for the log risk ratio. Row (*c*) shows the region of (β, γ) space in which the risk ratio is biased across the null hypothesis of $\beta = 0$. In the absence of bias, contour lines would be perfectly vertical and would match the value of β. In the absence of direction bias, heat maps for log[RR] would be entirely blue to the left of the $\beta = 0$ line, and red to the right of it. The greater the mismatch between β and the corresponding log[RR] contour line, the greater the bias. The bottom row shows the region of bias across the null, in which log[RR] and β have different signs. Study setup: clusters of size two with exactly one subject per cluster with $x = 1$. Values of the ω/α summarize the relative contribution of within-cluster transmission compared to transmission from the community. Observation time T was chosen such that cumulative incidence when $\beta = 0$ and $\gamma = 0$ is approximately 0.15 in all of the plots.

the hazard ratio, the risk ratio is always biased unless $\beta = 0$ and either $\omega = 0$ or $\gamma = 0$ holds. For all other combinations of parameters, whenever the risk ratio is not biased across the null, it is biased towards the null hypothesis of $\beta = 0$.

3.2. General clusters

The results derived above address performance of the risk ratio in clusters of size two. However, most empirical cluster cohort studies of infectious diseases involve variable cluster sizes and a more complex design. Several factors may influence the behaviour of the risk ratio in empirical studies, including epidemiologic features like the exogenous force of infection $\alpha_i(t)$, the force of contagion $\omega_{ikj}(t)$, and aspects of study design such as experimental assignment of the covariate x, the duration and variability of observation time T_i, cluster size distribution, or selection of clusters with or without infected individuals at baseline.

Result 3.5. (No within-cluster contagion.) Suppose $\omega_{ikj}(t) = 0$ for all t and $x_i = (x_{i1}, \ldots, x_{in_i})$ is independent of $\{\alpha_i(t), n_i, T_i\}$. Then the risk ratio is direction-unbiased.

Results 3.1 and 3.5 confirm the intuition that when there is no within-cluster contagion, and the covariate is

independent of the force of infection and observation time, the risk ratio is direction-unbiased.

Result 3.6. (Independent x.) Suppose the covariates $x_i = (x_{i1}, \ldots, x_{in_i})$ are jointly independent and x_i is independent of $\{\alpha_i(t), \omega_{ikj}(t), n_i, T_i\}$. Then the risk ratio is direction-unbiased.

Joint independence of within-cluster covariates guarantees direction-unbiasedness for any parameter values.

The risk ratio is not generally direction-unbiased when the joint distribution of x_i is dependent. For example, bias across the null may occur under two common randomization schemes used in clinical trials: 'block randomization' within clusters, when a fixed number of subjects per cluster have $x = 1$ with $\Pr(x_i) = \binom{n_i}{k_i}^{-1} 1\{\sum_j x_{ij} = k_i\}$, and 'cluster randomization' with $x_{ij} = 1$ for all j in some subset of clusters, and $x_{ij} = 0$ for all j in the remaining subset. In general, when the joint distribution of x_i is not independent, or when there is heterogeneity in n_i, $\alpha_i(t)$, or $\omega_{ikj}(t)$ across clusters, the risk ratio need not be direction-unbiased, even when $\gamma = 0$. Dependence in x_i may occur in observational studies, where x_i may be dependent due to shared environment, genetic factors or other forms of dependence within clusters. For example, an observational study of heterosexual partnerships might exhibit dependence on a covariate for gender.
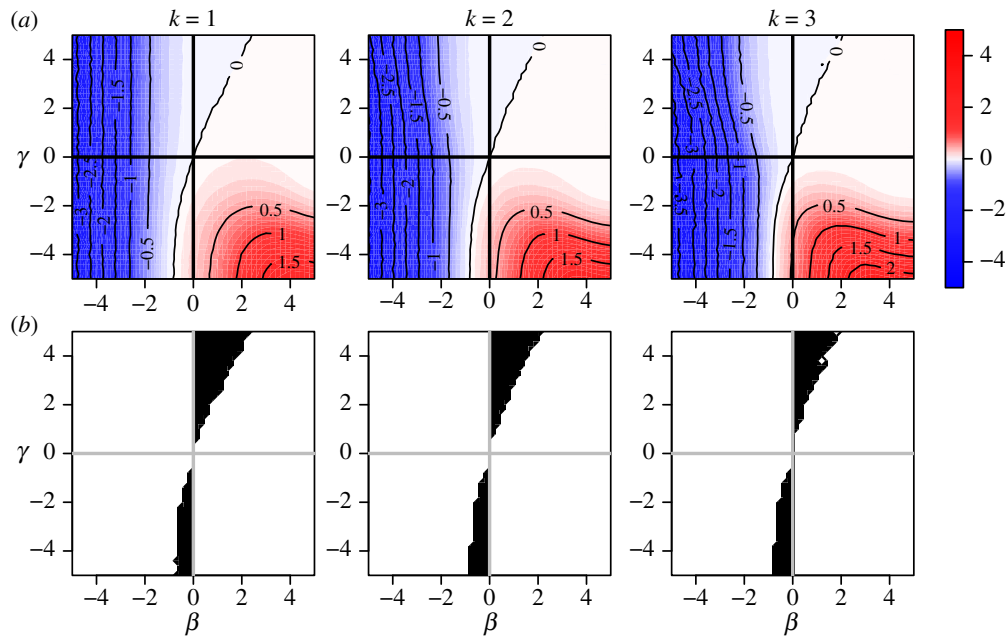
Figure 4. Risk ratio (log) and region of bias across the null as a function of true covariate effects β and γ when cluster size is constant ($n_i = 4$ for all i) and distribution of x is block randomized: $\sum_{j=1}^{n_i} x_{ij} = k$. Row (a) shows log[RR], and (b) shows the region of bias across the null (where log[RR] and β have opposite signs). In all plots $\alpha = 0.0001$, $\omega = 0.01$, $N = 500$, $T_i = 450$ and no subjects are infected at time zero.

3.3. Simulation results

Analytical expressions for the bias of the risk ratio as an approximation to the hazard ratio are intractable in general. However, simulations can provide further insight under particular epidemiologic and study design parameters. In simulations, we vary the distribution of covariates x_i, cluster size n_i, observation time T_i, infected cluster members at baseline and values of force of infection parameters α and ω, which are assumed to be constant over time and clusters. A comprehensive set of simulation results and R code [54] for replicating the simulations appear in the electronic supplementary material.

Some properties of the two person-cluster case hold in more complex scenarios. Figure 4 shows results for clusters of size four and block randomized distribution of x, such that each cluster has exactly k subjects with $x = 1$, $k = 1, 2,$ 3. The behaviour of the bias in figure 4 mimics that of the two-person cluster case. We demonstrated analytically in result 3.5 that direction-unbiasedness under no within-cluster contagion holds under independence of x_i and cluster level parameters $\alpha_i(t)$, n_i and T_i. The simulation shows that results under constant cluster size and block randomized x are similar to those in the two-person cluster case for sufficiently large observation times T_i.

It follows from result 3.6 that the risk ratio is direction-unbiased under independent Bernoulli assignment of x_i. In practical intervention trials, many studies in small clusters employ block or cluster randomization. Simulation results show that both of these methods result in bias across the null in a region of (β, γ) space. Figure 4 shows that block randomized distribution of x_i can result in bias across the null when β and γ have the same sign and γ is more extreme than β. Figure 5 illustrates cluster randomized distribution of x_i, showing bias across the null in regions where β and γ have opposite sign.

When cluster size n_i varies, bias patterns can change substantially with the nature of dependence in the distribution of x_i. Even under block randomization, the pattern of bias across the null differs depending on allocation proportion, and generally worsens with imbalance between $\Pr(x_{ij} = 1)$ and

$\Pr(x_{ij} = 0)$. Figure 6 illustrates bias across the null under variable cluster sizes with exactly one subject per cluster having $x = 1$, and figure 7 shows balanced block randomized x under variable cluster sizes. It is not necessary for γ to be more extreme than β, nor must these parameters have the same sign, to observe bias across the null. While regions where the risk ratio exhibits bias across the null become smaller in figure 7 compared to figure 6, in both cases it is still present when $\gamma = 0$. Thus, the desirable property of direction-unbiasedness under homogeneous infectiousness (result 3.3 in two-person cluster case) disappears when cluster sizes vary.

The duration of observation influences the size of the region in (β, γ) parameter space where the risk ratio exhibits direction bias. Longer observation times increase the region of direction bias under block randomized distribution of x and reduce the size of this region under cluster randomized distribution of x. When the distribution of x is jointly independent, the risk ratio is always direction unbiased; however, increasing the observation time increases the absolute value of the bias. When observation time varies from cluster to cluster, the behaviour of the bias remains similar to the case where all clusters are observed for the same duration, when all other study design parameters are kept the same.

In real-world cohort studies of infectious disease outcomes, researchers often select clusters (e.g. households) based on infection outcomes detected at baseline (sometimes called 'index' cases), especially for diseases with low overall prevalence or community force of infection, and risk ratios are computed at follow-up for cluster members susceptible at baseline. Simulation results, given in the electronic supplementary material, show that when subjects are infected at baseline, resulting direction bias depends on the distribution of x among infected and uninfected subjects at baseline.

4. Discussion

We have applied a standard and widely accepted measure of association to outcomes generated by a canonical
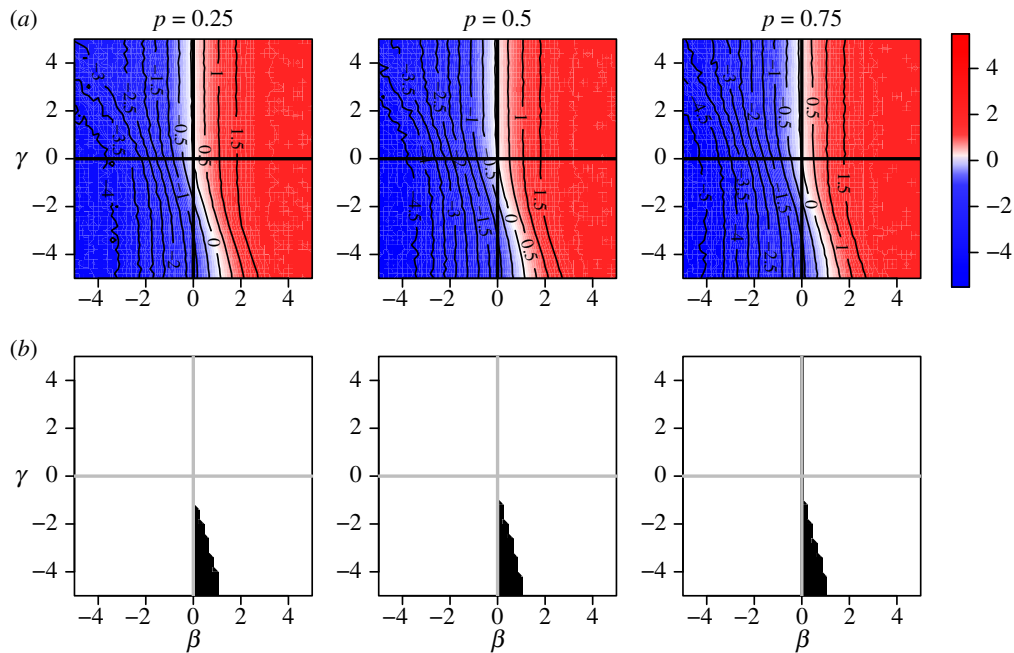
Figure 5. Risk ratio (log) and region of bias across the null as a function of true covariate effects β and γ when cluster size is constant and x is cluster randomized: proportion p of clusters have $\sum_{j=1}^{n_i} x_{ij} = 4$, and remaining $1 - p$ have $\sum_{j=1}^{n_i} x_{ij} = 0$. Row (a) shows the value of log[RR], and (b) shows the region of bias across the null. In all plots $\alpha = 0.0001$, $\omega = 0.01$, $N = 500$, $n_i = 4$, $T_i = 450$ and no subjects are infected at time zero.
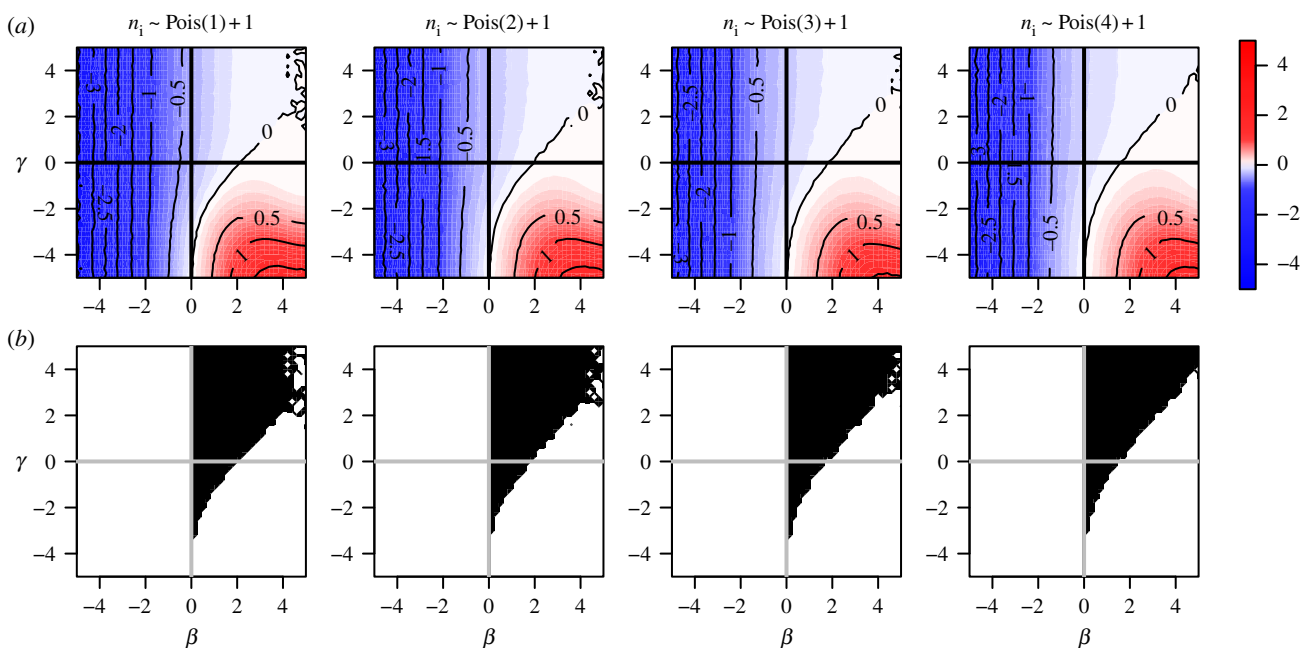

Figure 6. Risk ratio (log) and region of bias across the null as a function of true covariate effects β and γ when cluster size $n_i \sim \text{Pois}(\mu) + 1$ and x is block randomized such that $\sum_{j=1}^{n_i} x_{ij} = 1$ for all i. Row (a) shows log[RR], and (b) shows the region of bias across the null. In all plots $\alpha = 0.0001$, $\omega = 0.01$, $N = 500$, no subjects are infected at time zero and observation time T_i is chosen such that cumulative incidence when $\beta = 0$ and $\gamma = 0$ is approximately 0.15.

stochastic model of infectious disease contagion. Infectious disease epidemiologists have warned that simplistic summaries of association can be misleading in the presence of contagion, but none have explained formally how and why these summaries depend on the susceptibility and infectiousness effects of the covariate. The results presented here provide the first formal evidence of the pathological properties of the risk ratio under contagion. When the distribution of a covariate is dependent within clusters and associated with both susceptibility to infection and transmissibility once infected, the risk ratio for that covariate may imply an aggregate effect whose direction is opposite that of its individual-level effect on susceptibility to

infection. This form of grossly misleading error may occur even when the covariate effects on susceptibility (β) and infectiousness (γ) have the same sign. For example, a vaccine might protect individuals who receive it from infection ($\beta < 0$), and reduce their infectiousness ($\gamma < 0$) if they become infected. Clearly such a vaccine should be regarded as helpful, but the risk ratio measured from a block-randomized trial could show RR ≥ 1.

Our findings apply to many other traditional measures of association between a covariate and outcome under contagion. For example, the odds ratio (OR) always indicates the same effect direction as the risk ratio. That is, if RR < 1
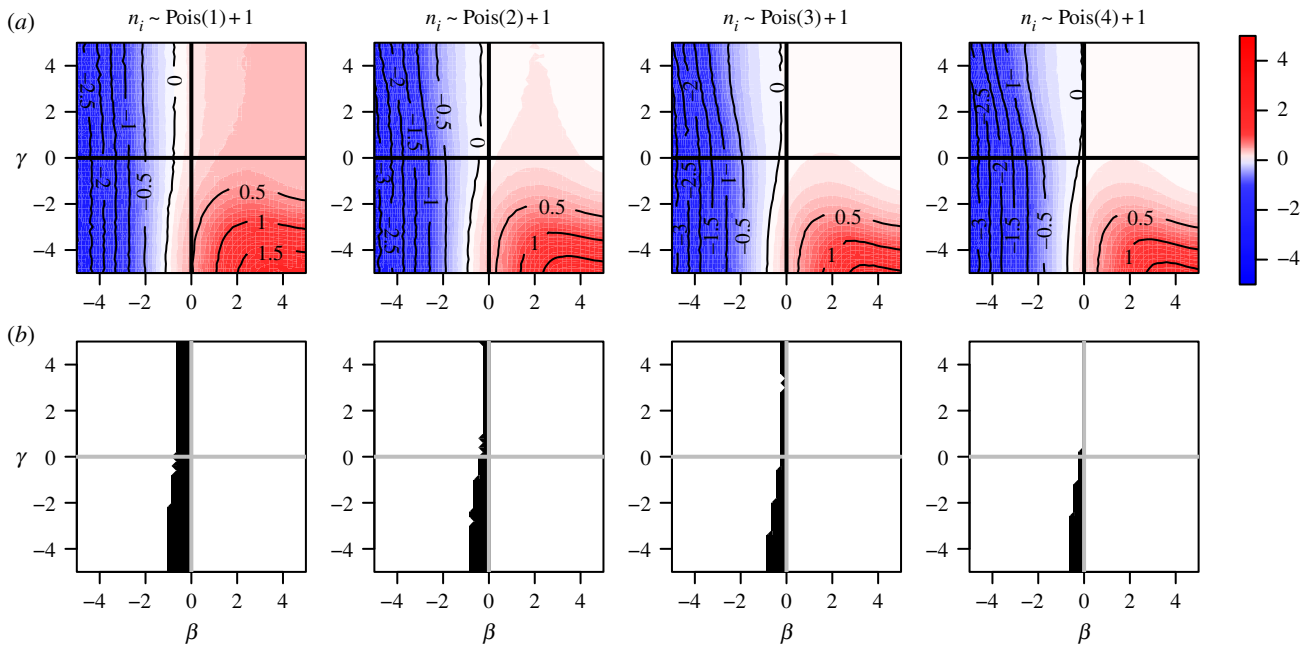
Figure 7. Risk ratio (log) and region of bias across the null as a function of true covariate effects β and γ when cluster size $n_i \sim \text{Pois}(\mu) + 1$ and x is block randomized such that $\sum_{j=1}^{n_i} x_{ij} = \lfloor n_i/2 \rfloor$ for all i. Row (a) shows log[RR], and (b) shows the region of direction bias. In all plots $\alpha = 0.0001$, $\omega = 0.01$, $N = 500$, no subjects are infected at time zero, and observation time T_i is chosen such that cumulative incidence when $\beta = 0$ and $\gamma = 0$ is approximately 0.15.

then OR < 1, and if RR > 1 then OR > 1. Therefore, the odds ratio can be expected to exhibit bias across the null under contagion whenever the risk ratio does. Likewise, the risk difference, attributable risk, and some measures of vaccine efficacy will suffer from direction bias under precisely the same conditions as the risk ratio.

Characterizing the bias of the risk ratio as an approximation to the susceptibility effect of a covariate gives clues about the performance of the risk ratio in a variety of empirical settings. Under the transmission model described by (2.4), the effect of any covariate x can be represented by a point in the two-dimensional (β, γ) space shown in figures 3–7 according to its true effect on susceptibility (β) and infectiousness (γ). In particular, when $x = 1$ denotes treatment or vaccination intended to reduce both susceptibility and infectiousness, we would expect its effect to be localized in the lower left quadrant of this space. For example, challenge studies have demonstrated the protective effect of influenza vaccine against infection ($\beta < 0$), and against infectiousness via reduction in viral shedding ($\gamma < 0$) [55]. We have shown that the risk ratio describing vaccine effect may be biased across the null in this setting (i.e. RR > 1 even though the vaccine protects against susceptibility and infectiousness). The likelihood of such egregious bias depends on disease parameters α and ω, and may be more likely to arise when using a block randomization study design. As another example, antiretroviral drugs may be given to people at risk for HIV infection as 'pre-exposure prophylaxis' to reduce susceptibility to HIV infection ($\beta < 0$) [56]. But antiretroviral HIV drugs also reduce viral load, and hence diminish infectiousness of people living with HIV ($\gamma < 0$) [57]. Susceptibility and infectiousness effect sizes of antiretrovirals are quite large, which could make it less likely that the risk ratio would exhibit direction bias in this setting.

When x is a variable associated with an increased risk of infection, it may be localized in the upper right quadrant of the (β, γ) covariate effect space. For example, a history of injection drug use is a risk factor for both acquiring ($\beta > 0$) and

transmitting ($\gamma > 0$) HIV and hepatitis C infections [58], and the risk ratio may be biased across the null under some study designs, particularly block randomization, in this setting. The region in the (β, γ) space where risk factors exhibit direction bias is generally larger than that for protective covariates conditional on the study design and duration of observation.

Finally, a covariate may have differential effects on susceptibility and infectiousness. It may, for instance, increase susceptibility to infection, but reduce infectiousness, placing it in the lower right quadrant of the (β, γ) space. For example, HIV infection may increase susceptibility to tuberculosis infection and clinical disease ($\beta > 0$) [59]. At the same time, studies have shown that people with tuberculosis co-infected with HIV may be less infectious than HIV-negative individuals ($\gamma < 0$) [60]. We have shown in simulation that when all subjects in the cluster have the same value of a covariate (e.g. under cluster randomization), the risk ratio may be biased across the null in the regions of the (β, γ) space where the signs of β and γ are opposite. Bias across the null also occurs under some block randomized designs when cluster sizes vary.

The bias of the risk ratio as an approximation to the hazard ratio (2.5) can be readily understood in terms of concepts already familiar to epidemiologists [9]. When the covariate x is dependent within clusters, other subjects' covariate values can be regarded as a common cause of both a given subject's covariate value (via dependence), and that subject's outcome (via contagion). Omitting or failing to condition on this common cause can result in bias. In other words, the relationship between an individual's covariate x and their infection outcome is subject to confounding [9,61], and bias across the null in this scenario is an example of Simpson's paradox [62–65].

The risk ratio is a valid statistical estimand: it summarizes the marginal association between x and infection. However, if investigators are interested in the *causal* direct (i.e. susceptibility) effect of treatment on the person who receives it [24,28], the risk ratio may give a very misleading estimate of this quantity. One striking consequence of result 3.4 is

that bias across the null can occur even when subjects are the same except for their value of x (exchangeable) and treatment (i.e. $x = 1$) is randomized and balanced within each cluster. The primary factor driving these results is contagion; bias across the null can occur even in the absence of unmodelled within-cluster heterogeneity, imbalance in covariate values, or heterogeneity in contact patterns [27].

Whether severe bias occurs in a particular empirical investigation depends on the epidemiologic features (i.e. $\alpha_i(t)$, $\omega_{ikj}(t)$) of the disease under study, the distribution of cluster size n_i and observation time T_i, and the distribution of x within clusters. Recently, researchers have called for more comprehensive simulation models to assist in the design of intervention studies for infectious disease outcomes [66]. Simulations such as those presented here may give insight into the anticipated properties of effect measures like the risk ratio under realistic models of infectious disease transmission. When a disease is only weakly contagious within clusters or when within-cluster transmissibility is less than the exogenous force of infection, bias across the null may be less likely to occur. This may be the situation in many cohort studies of infectious diseases. Based on result 3.6, researchers who wish to report the naive risk ratio may avoid direction bias when they can ensure that treatment assignment has independent Bernoulli distribution. The use of the risk ratio may also be justified in observational studies in which covariates of interest are independent or only weakly dependent within clusters. In a wide variety of empirical dependence settings in which infection is only weakly contagious, the risk ratio may be a reasonable estimator of the ratio of instantaneous risks. However, in studies of highly contagious outcomes (e.g. Ebola) in which ω/α is large and the protective effect of the intervention is weak, the risk ratio may exhibit direction bias under some designs.

Our analysis is subject to several limitations. First, we have focused primarily on the conditions under which the risk ratio can exhibit the most misleading form of bias—bias across the null. But even when direction-unbiasedness holds, the risk ratio is almost always a biased approximation of the covariate effect on susceptibility [27,28,31,34,67]. When the magnitude of an estimated risk ratio is important for decision-making, a more detailed study of bias may be warranted. Second, the transmission model in this analysis is represented by a standard stochastic susceptible–infective model with subject-specific covariates and an exogenous force of infection. This setting provides a simple generative model that incorporates features of infectious disease contagion relevant to the properties of the risk ratio. However, this model does not capture several important aspects of infectious disease dynamics, including recovery, removal, re-infection or multiple infections. Further investigation is necessary to determine properties of the risk ratio under more complex assumptions about disease dynamics.

Finally, the purpose of this work is to explain why the risk ratio may not be a satisfactory measure of association for epidemiologic studies of infectious disease outcomes. We have not addressed correction of the risk ratio, nor proposed alternative strategies for estimating the hazard ratio. Fortunately, epidemiologists have developed deterministic and stochastic models of infectious disease transmission in groups that take exposure to infection into account [37,43,44]. Several researchers have developed inferential approaches that capture infectious disease transmission dynamics and permit adjustment for individual-level factors [36,37,68–70]. Some analyses of contagious outcomes adjust for variables that may be correlated with exposure to infectiousness [71–74]. It remains an open question whether standard regression adjustment using a summary of infection outcomes of other individuals can deliver risk ratio estimates that are direction-unbiased.

References

1. Sinclair JC, Bracken MB. 1994 Clinically useful measures of effect in binary analyses of randomized trials. *J. Clin. Epidemiol.* **47**, 881–889. (doi:10.1016/0895-4356(94)90191-0)

2. Davies HTO, Crombie IK, Tavakoli M. 1998 When can odds ratios mislead? *Br. Med. J.* **316**, 989–991. (doi:10.1136/bmj.316.7136.989)

3. Bracken MB, Sinclair JC. 1999 Avoidable systematic error in estimating treatment effects must not be tolerated. *Br. Med. J.* **317**, 124.

4. Skov T, Deddens J, Petersen MR, Endahl L. 1998 Prevalence proportion ratios: estimation and hypothesis testing. *Int. J. Epidemiol.* **27**, 91–95. (doi:10.1093/ije/27.1.91)

5. Jewell N. 2003 *Statistics for epidemiology*. Boca Raton, FL: Taylor & Francis.

6. Greenland S. 2004 Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case–control studies. *Am. J. Epidemiol.* **160**, 301–305. (doi:10.1093/aje/kwh221)

7. Liberman AM. 2005 How much more likely? The implications of odds ratios for probabilities. *Am. J. Eval.* **26**, 253–266. (doi:10.1177/1098214005275825)

8. Katz KA. 2006 The (relative) risks of using odds ratios. *Arch. Dermatol.* **142**, 761–764. (doi:10.1001/archderm.142.6.761)

9. Rothman KJ, Greenland S, Lash TL. 2008 *Modern epidemiology*. Philadelphia, PA: Lippincott Williams & Wilkins.

10. Lumley T, Kronmal R, Ma S. 2006 Relative risk regression in medical research: models, contrasts, estimators and algorithms. Technical Report 293 UW Biostatistics Working Paper Series. See http://www.bepress.com/uwbiostat/paper293.

11. McNutt LA, Wu C, Xue X, Hafner JP. 2003 Estimating the relative risk in cohort studies and clinical trials

of common outcomes. *Am. J. Epidemiol.* **157**, 940–943. (doi:10.1093/aje/kwg074)

12. Zou G. 2004 A modified Poisson regression approach to prospective studies with binary data. *Am. J. Epidemiol.* **159**, 702–706. (doi:10.1093/aje/kwh090)

13. Spiegelman D, Hertzmark E. 2005 Easy SAS calculations for risk or prevalence ratios and differences. *Am. J. Epidemiol.* **162**, 199–200. (doi:10.1093/aje/kwi188)

14. Zou G, Donner A. 2013 Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat. Methods. Med. Res.* **22**, 661–670. (doi:10.1177/0962280211427759)

15. Jackson ML et al. 2011 Serologically confirmed household transmission of 2009 pandemic influenza A (H1N1) virus during the first pandemic wave—New York City, April–May 2009. *Clin. Infect. Dis.* **53**, 455–462. (doi:10.1093/cid/cir437)

16. Kim CY et al. 2012 Secondary household transmission of 2009 pandemic influenza A (H1N1) virus among an urban and rural population in Kenya, 2009–2010. *PLoS ONE* **7**, e38166. (doi:10.1371/journal.pone.0038166)

17. Bower H et al. 2016 Effects of mother's illness and breastfeeding on risk of Ebola virus disease in a cohort of very young children. *PLoS Negl. Trop. Dis.* **10**, e0004622. (doi:10.1371/journal.pntd.0004622)

18. Dowell SF, Mukunu R, Ksiazek TG, Khan AS, Rollin PE, Peters C. 1999 Transmission of Ebola hemorrhagic fever: a study of risk factors in family members, Kikwit, Democratic Republic of the Congo, 1995. *J. Infect. Dis.* **179**, S87–S91. (doi:10.1086/514284)

19. Araujo S, Rezende MMF, de Sousa DCR, Rosa MR, dos Santos DC, Goulart LR, Goulart IMB. 2015 Risk–benefit assessment of *Bacillus* Calmette-Guérin vaccination, anti-phenolic glycolipid I serology, and Mitsuda test response: 10-year follow-up of household contacts of leprosy patients. *Rev. Soc. Bras. Med. Trop.* **48**, 739–745. (doi:10.1590/0037-8682-0245-2015)

20. Seward JF, Zhang JX, Maupin TJ, Mascola L, Jumaan AO. 2004 Contagiousness of varicella in vaccinated cases: a household contact study. *J. Am. Med. Assoc.* **292**, 704–708. (doi:10.1001/jama.292.6.704)

21. Donnelly P. 1993 The correlation structure of epidemic models. *Math. Biosci.* **117**, 49–75. (doi:10.1016/0025-5564(93)90017-5)

22. Halloran ME, Struchiner CJ. 1991 Study designs for dependent happenings. *Epidemiology* **2**, 331–338. (doi:10.1097/00001648-199109000-00004)

23. Halloran ME, Haber M, Longini Jr IM. 1992 Interpretation and estimation of vaccine efficacy under heterogeneity. *Am. J. Epidemiol.* **136**, 328–343. (doi:10.1093/oxfordjournals.aje.a116498)

24. Halloran ME, Struchiner CJ. 1995 Causal inference in infectious diseases. *Epidemiology* **6**, 142–151. (doi:10.1097/00001648-199503000-00010)

25. Longini IM, Koopman JS, Monto AS, Fox JP. 1982 Estimating household and community transmission

parameters for influenza. *Am. J. Epidemiol.* **115**, 736–751. (doi:10.1093/oxfordjournals.aje.a113356)

26. Longini IM, Koopman JS, Haber M, Cotsonis GA. 1988 Statistical inference for infectious diseases risk—specific household and community transmission parameters. *Am. J. Epidemiol.* **128**, 845–859. (doi:10.1093/oxfordjournals.aje.a115038)

27. Koopman JS, Longini IM, Jacquez JA, Simon CP, Ostrow DG, Martin WR, Woodcock DM. 1991 Assessing risk factors for transmission of infection. *Am. J. Epidemiol.* **133**, 1199–1209. (doi:10.1093/oxfordjournals.aje.a115832)

28. Halloran ME, Struchiner CJ, Longini IM. 1997 Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am. J. Epidemiol.* **146**, 789–803. (doi:10.1093/oxfordjournals.aje.a009196)

29. Becker NG, Starczak DN. 1998 The effect of random vaccine response on the vaccination coverage required to prevent epidemics. *Math. Biosci.* **154**, 117–135. (doi:10.1016/S0025-5564(98)10048-2)

30. Chick SE, Barth-Jones DC, Koopman JS. 2001 Bias reduction for risk ratio and vaccine effect estimators. *Stat. Med.* **20**, 1609–1624. (doi:10.1002/sim.788)

31. Eisenberg JN, Lewis BL, Porco TC, Hubbard AH, Colford JM. 2003 Bias due to secondary transmission in estimation of attributable risk from intervention trials. *Epidemiology* **14**, 442–450. (doi:10.1097/01.ede.0000071411.19255.4c)

32. Koopman JS. 2004 Modeling infection transmission. *Annu. Rev. Public. Health.* **25**, 303–326. (doi:10.1146/annurev.publhealth.25.102802.124353)

33. Pitzer VE, Basta NE. 2012 Linking data and models: the importance of statistical analyses to inform models for the transmission dynamics of infections. *Epidemiology* **23**, 520–522. (doi:10.1097/EDE.0b013e31825902ab)

34. O'Hagan JJ, Lipsitch M, Hernán MA. 2014 Estimating the per-exposure effect of infectious disease interventions. *Epidemiology* **25**, 134–138. (doi:10.1097/EDE.0000000000000003)

35. Sharker Y, Kenah E. 2017 Estimation of the household secondary attack rate: binomial considered harmful. (http://arxiv.org/abs/1705.01135)

36. Kenah E. 2015 Semiparametric relative-risk regression for infectious disease transmission data. *J. Am. Stat. Assoc.* **110**, 313–325. (doi:10.1080/01621459.2014.896807)

37. Rampey AH, Longini IM, Haber M, Monto AS. 1992 A discrete-time model for the statistical analysis of infectious disease incidence data. *Biometrics* **48**, 117–128. (doi:10.2307/2532743)

38. Halloran ME, Longini IM, Haber MJ, Struchiner CJ, Brunet RC. 1994 Exposure efficacy and change in contact rates in evaluating prophylactic HIV vaccines in the field. *Stat. Med.* **13**, 357–377. (doi:10.1002/sim.4780130404)

39. Rhodes PH, Halloran ME, Longini IM. 1996 Counting process models for infectious disease data: distinguishing exposure to infection from susceptibility. *J. R. Stat. Soc.: Ser. B* **58**, 751–762.

40. Goldstein E, Pitzer JJ, O'hagan VE, Lipsitch M. 2017 Temporally varying relative risks for infectious diseases: implications for infectious disease control. *Epidemiology* **28**, 136–144. (doi:10.1097/EDE.0000000000000571)

41. Scott P, Herzog SA, Auranen K, Dagan R, Low N, Egger M, Heijne JC. 2014 Timing of bacterial carriage sampling in vaccine trials: a modelling study. *Epidemics* **9**, 8–17. (doi:10.1016/j.epidem.2014.08.003)

42. Greenland S, Frerichs RR. 1988 On measures and models for the effectiveness of vaccines and vaccination programmes. *Int. J. Epidemiol.* **17**, 456–463. (doi:10.1093/ije/17.2.456)

43. Anderson RM, May RM. 1992 *Infectious diseases of humans: dynamics and control.* New York, NY: Oxford University Press.

44. Andersson H, Britton T. 2000 *Stochastic epidemic models and their statistical analysis.* New York, NY: Springer-Verlag.

45. Kenah E. 2013 Non-parametric survival analysis of infectious disease data. *J. R. Stat. Soc.: Ser. B* **75**, 277–303. (doi:10.1111/j.1467-9868.2012.01042.x)

46. Smith P, Rodrigues L, Fine P. 1984 Assessment of the protective efficacy of vaccines against common diseases using case–control and cohort studies. *Int. J. Epidemiol.* **13**, 87–93. (doi:10.1093/ije/13.1.87)

47. VanderWeele TJ, Tchetgen Tchetgen EJ. 2011 Bounding the infectiousness effect in vaccine trials. *Epidemiology* **22**, 686–693. (doi:10.1097/EDE.0b013e31822708d5)

48. VanderWeele TJ, Tchetgen Tchetgen EJ, Halloran ME. 2012 Components of the indirect effect in vaccine trials: identification of contagion and infectiousness effects. *Epidemiology* **23**, 751–761. (doi:10.1097/EDE.0b013e31825fb7a0)

49. Ogburn EL, VanderWeele TJ. 2017 Vaccines, contagion, and social networks. *Ann. Appl. Stat.* **11**, 919–948. (doi:10.1214/17-AOAS1023)

50. Carpenter LM, Kamali A, Ruberantwari A, Malamba SS, Whitworth JA. 1999 Rates of HIV-1 transmission within marriage in rural Uganda in relation to the HIV sero-status of the partners. *Aids* **13**, 1083–1089. (doi:10.1097/00002030-199906180-00012)

51. Biraro S, Ruzagira E, Kamali A, Whitworth J, Grosskurth H, Weiss HA. 2013 HIV-1 transmission within marriage in rural Uganda: a longitudinal study. *PLoS ONE* **8**, e55060. (doi:10.1371/journal.pone.0055060)

52. Regev-Yochay G et al. 2009 Parental *Staphylococcus aureus* carriage is associated with staphylococcal carriage in young children. *Pediatr. Infect. Dis. J.* **28**, 960–965. (doi:10.1097/INF.0b013e3181a90883)

53. Leshem E et al. 2012 Transmission of *Staphylococcus aureus* from mothers to newborns. *Pediatr. Infect. Dis. J.* **31**, 360–363. (doi:10.1097/INF.0b013e318244020e)

54. R Core Team. 2017 *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org/.

55. Clements ML, Snyder MH, Sears SD, Maassab HF, Murphy BR. 1990 Evaluation of the infectivity, immunogenicity, and efficacy of live cold-adapted influenza B/Ann Arbor/1/86 reassortant virus vaccine in adult volunteers. *J. Infect. Dis.* **161**, 869–877. (doi:10.1093/infdis/161.5.869)

56. Fonner VA, Dalglish SL, Kennedy CE, Baggaley R, O'reilly KR, Koechlin FM, Rodolph M, Hodges-Mameletzis I, Grant RM. 2016 Effectiveness and safety of oral HIV preexposure prophylaxis for all populations. *AIDS (London, England)* **30**, 1973–1983. (doi:10.1097/QAD.0000000000001145)

57. Cohen MS *et al.* 2011 Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl. J. Med.* **365**, 493–505. (doi:10.1056/NEJMoa1105243)

58. Degenhardt L *et al.* 2016 Estimating the burden of disease attributable to injecting drug use as a risk factor for HIV, hepatitis C, and hepatitis B: findings from the global burden of disease study 2013. *Lancet Infect. Dis.* **16**, 1385–1398. (doi:10.1016/S1473-3099(16)30325-5)

59. Bell LC, Noursadeghi M. 2017 Pathogenesis of HIV-1 and *Mycobacterium tuberculosis* co-infection. *Nat. Rev. Microbiol*. (doi:10.1038/nrmicro.2017.128)

60. Cauthen GM, Dooley SW, Onorato IM, Ihle WW, Burr JM, Bigler WJ, Witte J, Castro KG. 1996 Transmission of *Mycobacterium tuberculosis* from tuberculosis patients with HIV infection or aids. *Am. J. Epidemiol.* **144**, 69–77. (doi:10.1164/ajrccm.164.12.2103078)

61. Greenland S, Morgenstern H. 1989 Ecological bias, confounding, and effect modification. *Int. J. Epidemiol.* **18**, 269–274. (doi:10.1093/ije/18.1.269)

62. Greenland S, Robins JM, Pearl J. 1999 Confounding and collapsibility in causal inference. *Stat. Sci.* **14**, 29–46. (doi:10.1214/ss/1009211805)

63. Arah OA. 2008 The role of causal reasoning in understanding Simpson's paradox, Lord's paradox, and the suppression effect: covariate selection in the analysis of observational studies. *Emerg. Themes. Epidemiol.* **5**, 1–5. (doi:10.1186/1742-7622-5-1)

64. Pearl J. 2000 *Causality*. New York, NY: Cambridge University Press.

65. Pearl J. 2014 Comment: understanding Simpson's paradox. *Am. Stat.* **68**, 8–13. (doi:10.1080/00031305.2014.876829)

66. Halloran ME *et al.*2017 Simulations for designing and interpreting intervention trials in infectious diseases. bioRxiv preprint (https://doi.org/10.1101/198051).

67. Staples PC, Ogburn EL, Onnela JP. 2015 Incorporating contact network structure in cluster randomized trials. *Sci. Rep.* **5**, 17581. (doi:10.1038/srep17581)

68. Longini IM, Koopman JS. 1982 Household and community transmission parameters from final distributions of infections in households. *Biometrics* **38**, 115–126. (doi:10.2307/2530294)

69. Haber M, Longini IM, Cotsonis GA. 1988 Models for the statistical analysis of infectious disease data. *Biometrics* **44**, 163–173. (doi:10.2307/2531904)

70. Becker NG, Britton T, O'Neill PD. 2003 Estimating vaccine effects on transmission of infection from household outbreak data. *Biometrics* **59**, 467–475. (doi:10.1111/1541-0420.00056)

71. Fine PE, Steme J, Pönnighaus J, Bliss L, Saul J, Chihana A, Munthali M, Wamdorff D. 1997 Household and dwelling contact as risk factors for leprosy in northern Malawi. *Am. J. Epidemiol.* **146**, 91–102. (doi:10.1093/oxfordjournals.aje.a009195)

72. Huang CC *et al.* 2014 The effect of HIV-related immunosuppression on the risk of tuberculosis transmission to household contacts. *Clin. Infect. Dis.* **58**, 765–774. (doi:10.1093/cid/cit948)

73. Martinez L, Sekandi JN, Castellanos ME, Zalwango S, Whalen CC. 2016 Infectiousness of HIV-seropositive patients with tuberculosis in a high-burden African setting. *Am. J. Respir. Crit. Care. Med.* **194**, 1152–1163. (doi:10.1164/rccm.201511-2146OC)

74. Staples PC, Prague M, Victor DG, Onnela JP 2016 Leveraging contact network information in clustered randomized trials of infectious processes. (http://arxiv.org/abs/1610.00039)