Heterogeneity in and correlation between host transmissibility and susceptibility can greatly impact epidemic dynamics

Beth M. Tuschhoff[a], David A. Kennedy[a,*]

[a] *Department of Biology, The Pennsylvania State University, University Park, Pennsylvania, United States of America*

Abstract

While it is well established that host heterogeneity in transmission and host heterogeneity in susceptibility each individually impact disease dynamics in characteristic ways, it is generally unknown how disease dynamics are impacted when both types of heterogeneity are simultaneously present. Here we explore this question. We first conducted a systematic review of published studies from which we determined that the effects of correlations have been drastically understudied. We then filled in the knowledge gaps by developing and analyzing a stochastic, individual-based SIR model that includes both heterogeneity in transmission and susceptibility and flexibly allows for positive or negative correlations between transmissibility and susceptibility. We found that in comparison to the uncorrelated case, positive correlations result in major epidemics that are larger, faster, and more likely, whereas negative correlations result in major epidemics that are smaller and less likely. We additionally found that, counter to the conventional wisdom that heterogeneity in susceptibility always reduces outbreak size, heterogeneity in susceptibility can lead to major epidemics that are larger and more likely than the homogeneous case when correlations between transmissibility and susceptibility are positive, but this effect only arises at small to moderate R_0. Moreover, positive correlations can frequently lead to major epidemics with subcritical R_0. Ultimately, we show that correlations between transmissibility and susceptibility profoundly impact disease dynamics.

Highlights

- Systematic review finds that effects of correlations on epidemics are understudied
- Positive correlations lead to larger, faster, more likely epidemics
- Negative correlations lead to smaller, less likely epidemics
- Positive correlations consistently lead to major epidemics with subcritical R_0

Keywords: infectious disease dynamics, heterogeneity in transmission, heterogeneity in susceptibility, correlation

1. Introduction

Some epidemics are fast and explosive while others are slow and meandering. These differences could arise in many ways, including from differences in factors such as R_0 values or transmission rates (Anderson and May, 1991; Heffernan et al., 2005). Here we explore whether heterogeneity in and correlations between host transmissibility and susceptibility can also drive changes in disease dynamics.

It is well established that host heterogeneity can have large impacts on disease dynamics (Dwyer et al., 1997; Lloyd-Smith et al., 2005; Gomes et al., 2014; Langwig et al., 2017; Gomes et al., 2022). Compared to the case without heterogeneity, heterogeneity in individuals' likelihoods of transmitting a pathogen once infected, hereafter referred to as "heterogeneity in transmission", results in epidemics that are rarer and more likely to go extinct but initially more explosive when they do take off (Lloyd-Smith et al., 2005). Likewise, heterogeneity in individuals' likelihoods of becoming infected, hereafter referred to as "heterogeneity in susceptibility", results in a lower peak number of cases and a smaller final epidemic size (Aguas et al.,

*Corresponding author: dak30@psu.edu

2020; Gomes et al., 2022; Montalbán et al., 2022). But what happens when a population has heterogeneity in both transmission and susceptibility? Moreover, how do individual-level correlations in transmissibility and susceptibility affect disease dynamics? Here we explore the effects of positive and negative correlations on disease dynamics, first by conducting a systematic literature review and second by developing a finite population model that flexibly allows for positive or negative correlations between host transmissibility and host susceptibility.

Correlations between transmissibility and susceptibility may be the rule rather than the exception. Positive correlations may arise in several ways. One way is that individuals may be heterogeneous in their contact rates. In such cases, it would be reasonable to expect individuals with more contacts to be at higher risk for infection and more likely to transmit a pathogen once infected, as might be the case for people with careers that require substantial public interface and social interactions (Baker et al., 2020). Another way that individuals can vary is in their risk taking behavior. Risky behaviors can make some individuals both more susceptible to infection and more transmissible given infection. For instance, individuals who engage in high risk sexual behaviors are more likely to both contract and transmit sexually transmitted infections (Ericksen and Trocki, 1992). Coinfection may also lead to positive correlations betweeen susceptibility and transmissibility when primary infections weaken immune responses (Galvin and Cohen, 2004; Zélé et al., 2018; Karvonen et al., 2019). Likewise, individuals that are immunocompromised for other reasons may be at greater risk for infection and have increased shedding, leading to more transmission (Permar et al., 2006; Memoli et al., 2014).

Negative correlations may also arise under biologically reasonable scenarios. For example, when severe disease results in an individual changing their behavior such that they reduce their contact rates, immunocompromised individuals would likely be at greater risk of infection but less likely to transmit infection once it occurs (Collins et al., 2020; Li et al., 2021). Similarly, if individuals that are resistant to infection are also more likely than average to develop asymptomatic infections, they may end up transmitting at higher rates than average because they are unaware of their infection status (Kalajdzievska and Li, 2011; Medzhitov et al., 2012). Certain types of public health interventions could additionally induce negative correlations between susceptibility and transmissibility. For instance, vaccination targeted towards people with highly social jobs could lower their overall susceptibility to infection, but if infected, these individuals may still transmit at high rates (Nunner et al., 2022; Park et al., 2022). Similarly, during disease outbreaks, front line workers may be at high risk for infection due to high rates of pathogen exposure, but may also be closely monitored and quarantined at the first signs of illness, reducing opportunities for transmission (Bielicki et al., 2020).

There has been substantial work conducted on understanding the role of host heterogeneity in transmission or susceptibility on disease dynamics. In this paper, we explored the role of interactions between these factors. We first conducted a systematic review of published studies that investigated the effects of correlations between transmissibility and susceptibility on disease dynamics to identify knowledge gaps in this topic area. To fill these knowledge gaps, we then developed a stochastic, individual-based SIR model that includes both heterogeneity in transmission and susceptibility and that flexibly allows for positive, negative, or zero correlations between transmissibility and susceptibility.

2. Definitions of the basic and effective reproductive numbers

The basic reproductive number R_0 is a valuable parameter for inferring disease dynamics and responding to infectious disease threats (Shaw and Kennedy, 2021). Using its standard definition, R_0 is the expected number of new infections caused by an initial infection in a fully susceptible population (Anderson and May, 1991). Nevertheless, despite this seemingly clear definition, there is ambiguity in its meaning for populations that are heterogeneous when transmissibility and susceptibility are correlated. This is because the value of R_0 in such situations depends on which individual is the initially infected one. To illustrate, consider that with heterogeneity in transmission and susceptibility, each individual $i = 1, ..., N$ has force of infection λ_i, risk of infection r_i, and probability of being the first infected p_i. Thus,

$$R_0 = \frac{\bar{\beta} \sum_i \sum_{j \neq i} p_i \lambda_i r_j}{\gamma}, \tag{1}$$

where $\bar{\beta}$ is a scaling constant for the average transmission rate of an infected individual and the denominator is the summed rates at which infection ends, illustrated here through a recovery rate γ. Note that $\sum_{j \neq i} r_j = \bar{r}N - r_i$ where $\bar{r}$ is the mean risk, so

$$R_0 = \frac{\bar{\beta} \sum_i p_i \lambda_i (\bar{r}N - r_i)}{\gamma} = \frac{\bar{\beta} \sum_i (p_i \lambda_i \bar{r}N - p_i \lambda_i r_i)}{\gamma}. \tag{2}$$

Under the assumption that $\bar{r}N$ is large in relation to the maximum r_i, which will be true at large N,

$$R_0 \approx \frac{\bar{\beta}\bar{r}N \sum_i p_i \lambda_i}{\gamma}. \tag{3}$$

When every individual is equally likely to be the first infected individual, $p_i = \frac{1}{N}$ for all i and R_0 can be defined as the product of mean transmissibility and mean susceptibility such that

$$R_0 \approx \frac{\bar{\beta}\bar{r}N \sum_i \frac{1}{N} \lambda_i}{\gamma} = \frac{\bar{\beta}\bar{\lambda}\bar{r}N}{\gamma}. \tag{4}$$

However, some individuals may be more likely to be the first infected individual than others. If the likelihood of being the first infection were exactly proportional to an individual's susceptibility, $p_i = \frac{r_i}{\sum_j r_j} = \frac{r_i}{\bar{r}N}$ and R_0 would be defined using the mean product of transmissibility and susceptibility,

$$R_0 \approx \frac{\bar{\beta}\bar{r}N \sum_i \frac{r_i}{\bar{r}N} \lambda_i}{\gamma} = \frac{\bar{\beta} \sum_i \lambda_i r_i}{\gamma}. \tag{5}$$

These values of Eqs. 4 and 5 need not be similar. Moreover, numerous other definitions of R_0 are possible, and depend on the distribution of p_i. In this manuscript, we used the definition of R_0 where every individual is equally likely to be the first infected (Eq 4).

The effective reproductive number R_e is the expected number of new infections caused by an infectious individual in a population composed of both susceptible and not susceptible individuals. Unlike R_0, R_e does not have an ambiguous meaning and so avoids the same definitional issues. This is because once an epidemic begins, individuals' chances of being infected will depend on their susceptibility. Hence, with S_t susceptible individuals and I_t infected individuals at time t, R_e can be defined as

$$R_e = \frac{\bar{\beta} \sum_{i=1}^{I_t} \sum_{j=1}^{S_t} \lambda_i r_j}{\gamma I_t}. \tag{6}$$

3. Systematic review

To explore the work that has previously been published on the effects of correlations between transmissibility and susceptibility on disease dynamics, we conducted a systematic review.

3.1. Methods

Our systematic review was conducted following the PRISMA guidelines (Fig 1; Page et al. 2021). We searched Google Scholar on January 30, 2024 using the following keywords:

("heterogeneity in susceptibility" OR "variation in susceptibility" OR "heterogeneous susceptibility" OR "differential susceptibility") AND ("heterogeneity in transmission" OR "heterogeneity in infectivity" OR "heterogeneous infectivity" OR "variation in infectivity" OR "variation in infectiousness" OR "differential infectivity") AND (correlation) AND (infectious disease).

We screened an initial 330 published papers by implementing the following selection criteria:

i The study must assess a mathematical model.

ii The model must include both heterogeneity in transmission, in terms of an individual's likelihood of spreading infection, and heterogeneity in susceptibility, in terms of an individual's likelihood of becoming infected.

iii The model must explore the effect of a correlation between transmissibility and susceptibility on infection dynamics.

iv The model must not implicitly assume a perfect positive correlation (e.g., an undirected network model).

A total of 9 papers fulfilled the selection criteria (Cardell and Kanouse, 1989; Becker and Marschner, 1990; Daley et al., 2000; Andreasen, 2011; Clancy and Pearce, 2013; Hickson and Roberts, 2014; Tkachenko et al., 2021; Kawagoe et al., 2021; Allard et al., 2023). From these selected papers, we recorded the author affiliations, the model structure, the assumptions, and what the model found with regard to the effects of positive or negative correlations between susceptibility and transmissibility on the probability of a major epidemic, the peak size, the peak time, the final epidemic size, and the time to the jth infection. For two of the papers, two different sets of assumptions were explored in each. We thus included each of these model versions as separate models in our analyses. This gave us a total of 11 models across the 9 papers that explored how heterogeneity in susceptibility and transmission affect disease dynamics in situations where susceptibility and transmissibility are positively or negatively correlated.

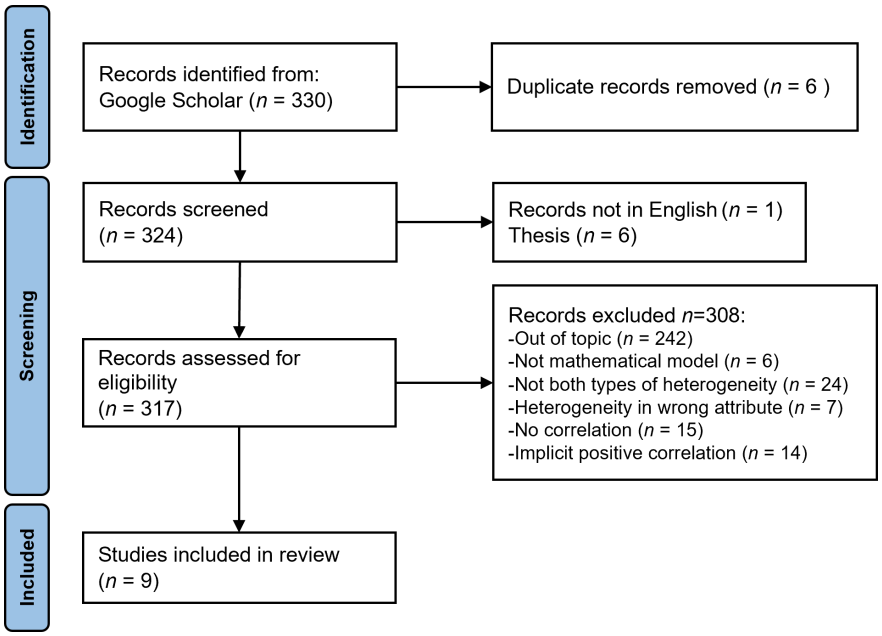

Figure 1: PRISMA systematic review framework.

3.2. Results

The 11 models from our 9 selected studies were built using various model structures and assumptions as presented in Fig 2. Eight of 11 models used an SIR-type model structure, 7 of 11 models were deterministic, and 4 of the 11 models assumed that host variation fell into discrete categories. The specific types of distributions used to model heterogeneity differed between models including gamma distributions, unnamed distributions where the authors set their own values, other distributions including lognormal or uniform, and a combination of gamma and another distribution. Ten out of 11 models assumed that the population was well-mixed with the exception being a directed network model. Ten out of 11 models also assumed that the size of the susceptible population changed over time due to individuals transitioning from susceptible to infectious, with the exception being a branching process model. When exploring the effects of correlations betweeen

susceptibility and transmissibility on disease dynamics, 4 of the 11 models kept the mean susceptibility and mean transmissibility constant (i.e., our preferred definition of R_0: Eq 4) and 5 of 11 kept the mean product of susceptibility and transmissibility constant (i.e., an alternative definition of R_0: Eq 5). The remaining 2 models ("Other") respectively kept the median product of susceptibility and transmissibility constant or kept the growth rate for the first month of the epidemic constant. We additionally found that the papers' authors came from various disciplines but were primarily associated with departments in the fields of physics, math, and biophysics (Fig 3).

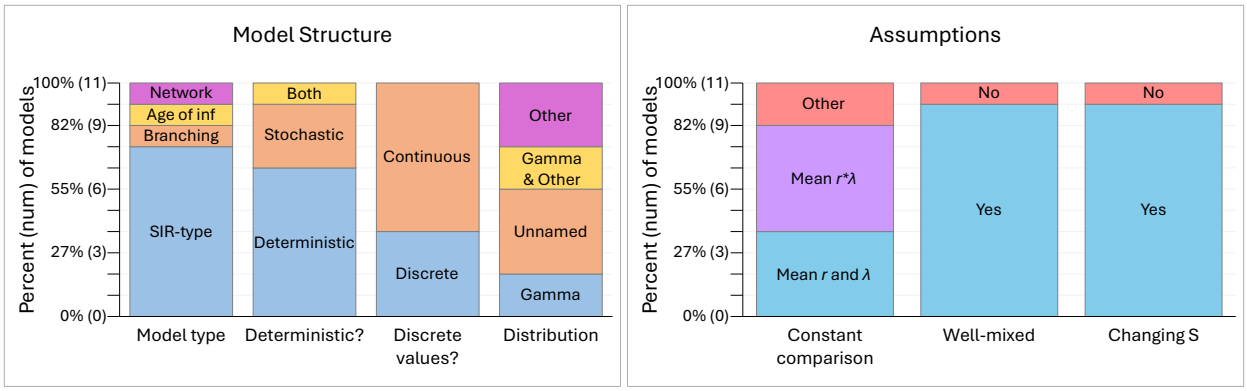

Figure 2: Models were built using various model structures and assumptions. These plots show the number of models that were built under each structure and assumption. For model structure, the model type is either an SIR-type (SIR, SI, etc.), a branching process, an age of infection model, or a directed network. Models were set up as deterministic, stochastic, or both to explore different results. The distributions used to model the values for susceptibility and transmissibility were either discrete or continuous and the type of distribution explored in each study was a gamma distribution, an unnamed distribution where authors set their own values, a gamma distribution plus another distribution, or solely another type of distribution like a lognormal or uniform distribution. For assumptions, models compare the effects of correlation to the uncorrelated and homogeneous cases by keeping the mean susceptibility and mean transmissibility constant (i.e., R_0 defined as in Eq 4; "Mean r and λ"), keeping the mean product of susceptibility and transmissibility constant (i.e., R_0 defined as in Eq 5; "Mean $r * \lambda$"), or by keeping the growth rate for the first month of the epidemic or the median product of susceptibility and transmissibility constant ("Other"). Most models assume a well-mixed population and a changing susceptible population size over time.

Among the 11 models, the most studied attributes were the final epidemic size and the peak time (Fig 4). The probability of a major epidemic, peak size, and time to the jth infection were not as commonly studied. These attributes were only explored by at most two models, and often the models disagreed on the results. In addition, the majority of models explored the effects of a positive correlation between susceptibility and transmissibility but not the effects of a negative correlation. Our review further showed that almost all of the models studied the effects of correlations relative to the case where populations were homogeneous rather than where populations were heterogeneous but susceptibility and transmissibility were uncorrelated. The lone exceptions were two models that investigated the impact of positive correlations on the final epidemic size. Thus, no models explored how negative correlations affected disease dynamics relative to populations with the same heterogeneity, but in which susceptibility and transmissibility were uncorrelated.

While most models seemed to agree that positive correlations lead to a smaller final epidemic size and earlier peak time than dynamics with homogeneous transmission and susceptibility, we did not find total agreement. Two of 2 models agreed that positive correlations result in larger final epidemic sizes compared to the no correlation case, and 2 of 2 models agreed that negative correlations lead to a smaller probability of a major epidemic in comparison to the homogeneous case. Nevertheless, we found substantial disagreement among many of the models in how correlations impact other epidemic features relative to the homogeneous case, including whether either type of correlation leads to a smaller or larger peak size, whether positive correlations lead to a smaller or larger probability of a major epidemic, whether negative correlations delay the timing of the peak, and whether negative correlations lead to a smaller, larger, or similar final epidemic size. We suspect the disagreement between models is due to differing model structures, different assumptions, different levels of heterogeneity, and different values of R_0. For example, the model that found that positive correlations result in larger final epidemic sizes kept the median of individuals' R_0 values constant for comparison between the heterogeneous and homogeneous cases, whereas all of the models that kept R_0

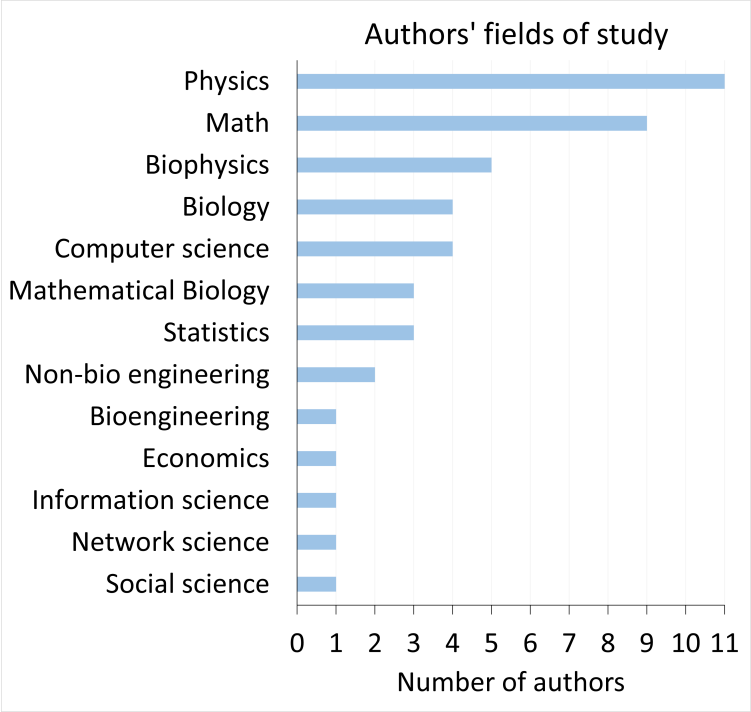

Figure 3: Authors of the selected studies in our systematic review tend to be associated with quantitative fields, especially physics, math, and biophysics. This plot shows the number of authors affiliated with each of the different fields of study represented in the selected studies. Authors' affiliations were identified by their department or research group and focus at the time the study was published. While there are 31 authors across the studies, some authors are represented multiple times as they are affiliated with multiple fields.

constant, regardless of how they defined R_0 (Eq 4 or 5), did not find this effect. Likewise, two models found that compared to the homogeneous case, the effects of a negative correlation on final epidemic size depended on the levels of heterogeneity.

Taken together, our systematic review reveals that the effects of correlation on disease dynamics were understudied by previous models. In particular, the effects of negative correlations and the comparison of dynamics to the uncorrelated case need to be studied. Additionally, the specific attributes of the probability of a major epidemic, peak size, and time to the jth infection were not as well-studied and should be examined further. Tables containing the full results of the systematic review as associated with each model are in the Supplement (Tables S1 and S2).

4. Model

To fill in the knowledge gaps and ambiguities identified from our systematic review, we generated an individual-based, stochastic SIR model that flexibly allows us to generate populations of hosts that are heterogeneous or not, and in which we can flexibly alter heterogeneity without altering R_0 (Eq 4). Using our model, we generated populations of hosts that have positive, negative or noncorrelated relationships between transmissibility and susceptibility. With this model, we explore the impact that positive and negative correlations have relative to the uncorrelated case and the homogeneous case. We assessed the same epidemic measures as in the systematic review as well as the effective reproductive number R_e.

4.1. Methods

For our model, we start with the basic SIR framework with ordinary differential equations

$$\frac{dS}{dt} = -\beta SI \tag{7}$$

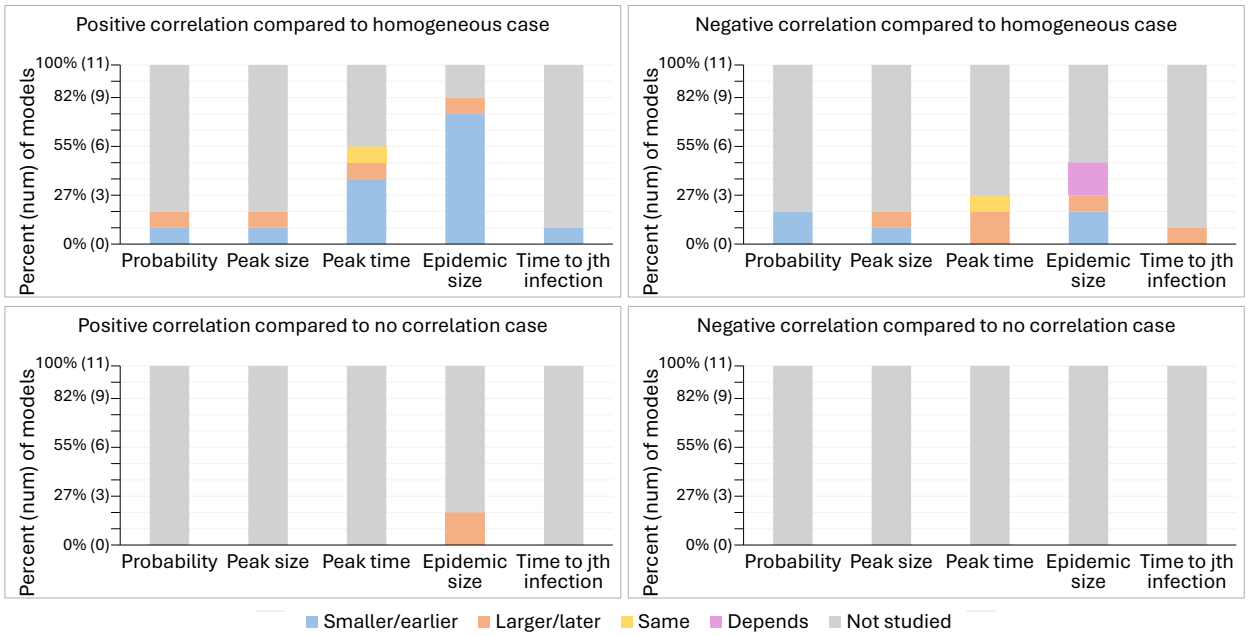

Figure 4: Models tend to disagree about the effects of correlation, and many epidemic measures are understudied, especially comparisons to the no correlation case. These plots show the effects of positive and negative correlations between heterogeneity in transmission and susceptibility on the probability of a major epidemic, peak size, peak time, final epidemic size, and time to the jth infection in comparison to disease dynamics with homogeneity or no correlation according to 11 models from the 9 studies included in the systematic review. The effect of correlation is classified for each model and measure as resulting in an attribute that is smaller/earlier (blue), larger/later (orange), the same (yellow), dependent on the levels of heterogeneity (purple), or not studied (gray) in comparison to the homogeneous or no correlation case.

$$\frac{dI}{dt} = \beta SI - \gamma I, \tag{8}$$

where β is the transmission rate and γ is the recovery rate. We then include heterogeneity in transmission and susceptibility. To do so, we first assume that each susceptible individual i has a unique susceptibility risk r_i, and each infected individual j has a unique force of infection λ_j, where prior to the initiation of an outbreak the expected values $\bar{r}$ and $\bar{\lambda}$ are both normalized to be 1 as described below. Next, we assume that the rate of transmission is multiplicatively determined by these parameters such that the transmission rate from individual i to j is $\beta_{i,j} = \bar{\beta}\lambda_i r_j$, where $\bar{\beta}$ is a constant that scales the average transmission rate, which we use to set R_0 according to Eq. 4.

To run our model, we assign each individual i, $i = 1, ..., N$ in the population a relative force of infection λ_i and a relative risk of being infected r_i. We set correlation ρ between these parameters using a normal copula with marginal gamma distributions where $\lambda_i \sim \text{Gamma}(m, 1/m)$ and $r_i \sim \text{Gamma}(k, 1/k)$. We chose to use gamma distributions because they are flexible and have been used to model heterogeneous populations previously (Dwyer et al., 1997; Langwig et al., 2017). Note that we facilitated comparisons across different levels of heterogeneity by setting the mean risk $\bar{r} = 1$ and mean force of infection $\bar{\lambda} = 1$, keeping R_0 constant, according to our working definition of R_0 (Eq 4). We describe the level of heterogeneity in transmission as C_t, which is the coefficient of variation in the force of infection across all potential hosts in the population, where the coefficient of variation is defined as the standard deviation divided by the mean. Being gamma distributed, this equates to $C_t = \frac{1}{\sqrt{m}}$. Similarly, we describe the level of heterogeneity in susceptibility using the coefficient of variation $C_s = \frac{1}{\sqrt{k}}$.

At the start of each simulation, we set parameters dictating the levels of heterogeneity present in the population (C_t, C_s), the correlation between transmissibility and susceptibility (ρ), the average transmission rate ($\bar{\beta}$), the recovery rate (γ), and the number of susceptible individuals and infected individuals at the start of the epidemic (S_0 and I_0). For our simulations, we set $C_t = 0, 0.5, 1,$ or 3, $C_s = 0, 0.5, 1,$ or 3,

$\rho = -1, -0.5, 0, 0.5$, or 1, $\bar{\beta} = 0.00008, 0.00011$, or 0.0003, $\gamma = 0.1$, $S_0 = 999$, and $I_0 = 1$, giving a population size of $S_0 + I_0 = 1,000$. We set $\bar{\beta}$ to correspond to R_0 values of $0.8, 1.1$, or 3, under the definition of R_0 in Eq. 4. We then randomly selected I_0 individuals to start as infected, where each individual has the same probability of being selected.

We simulated the epidemic using a Gillespie algorithm. In this algorithm, we determine the time to the next event, which is when an infected individual either infects a susceptible individual or recovers. Infection occurs at rate $\sum_{i=1}^{I}\sum_{j=1}^{S}\bar{\beta}\lambda_i r_j$, and recovery occurs at rate γI. If the next event is a new infection, we select a susceptible individual to become infected and move to the I class based on each individual's relative risk such that the probability of selecting individual i is equal to $\frac{r_i}{\sum_{j=1}^{S_t} r_j}$. If the next event is a host recovery, we randomly select an infected individual to recover where each individual has an identical likelihood of being chosen. At each time point, we record the number of S, I, and R individuals as well as the effective reproductive number R_e.

For each parameter combination of C_t, C_s, ρ, and R_0, we ran 200 epidemic simulations. After simulating the epidemics, we calculated the following epidemic features for each simulation in which there was a major epidemic:

1. The probability of a major epidemic, defined as the fraction of simulations in which the final epidemic size is greater than a set threshold,
2. The peak size, defined as the maximum number of I individuals at any point in the epidemic,
3. The peak time, defined as the time at which the peak size occurs,
4. The final epidemic size, defined as the difference between the population size and the number of S individuals at the end of the epidemic $N - S_{\text{final}}$,
5. The time to the jth infection, and
6. The effective reproductive number R_e defined at each time step t as in equation 6.

In practice, the threshold for calling something a major epidemic was determined by plotting a histogram of the final epidemic sizes and visually inspected the histogram for the place where there was a clear division between epidemics in which few (minor) or many (major) individuals became infected (Southall et al., 2023). Because the level of heterogeneity in susceptibility impacts epidemic size (Gomes et al., 2022; Montalbán et al., 2022), we set a different threshold for different values of C_s. Our threshold number of cases was 200 for $C_s = 0.5$, 100 for $C_s = 1$, and 50 for $C_s = 3$.

For each measure, we determined the mean or median to compare the measures across parameter combinations. Since the time to the jth infection is a time series with stochastic times, we first binned the output into 0.1 step size intervals to compare across simulations. To do so, we determined the status of this measure at each 0.1 step in time (i.e., at time 0, 0.1, 0.2,...) before calculating the median across the simulations at each time point over the epidemic.

4.2. Results

In Fig 5, we show the trajectory of epidemics for $R_0 = 3$ with different levels of heterogeneity in susceptibility, heterogeneity in transmission, and correlations between susceptibility and transmissibility, and we compare these trajectories to the case without heterogeneity and the case where susceptibility and transmissibility are uncorrelated. We found that correlations between transmissibility and susceptibility can greatly affect the probability of a major epidemic, peak size, peak time, final epidemic size, time to the jth infection, and R_e. We also explored how different R_0 may affect our results, using Eq. 4 for our definition of R_0. Fig 6 shows epidemic trajectories for $R_0 = 0.8$ and $R_0 = 1.1$ with high levels of heterogeneity ($C_s = C_t = 3$). As expected, as R_0 decreased, the probability of a major epidemic, peak size, and final epidemic size decreased. Also, the peak time and time to the jth infection tended to be later as R_0 decreased. In comparison to the uncorrelated case, with different R_0, we found the same general trends that we discuss below at varying degrees of emphasis for each epidemic measure. However, in comparison to the homogeneous case, positive correlations can lead to opposite effects for some measures when R_0 is large versus small (Supplement sections 3 and 4). In the remainder of the text, we focus primarily on presenting results for the case where $R_0 = 3$, but we specify when the results are different for smaller R_0.

We first examined the effect of heterogeneity alone on disease dynamics. As expected based on previous studies, we found that with heterogeneity in susceptibility alone ($C_s = 0.5, 1$, or 3 and $C_t = 0$), epidemics

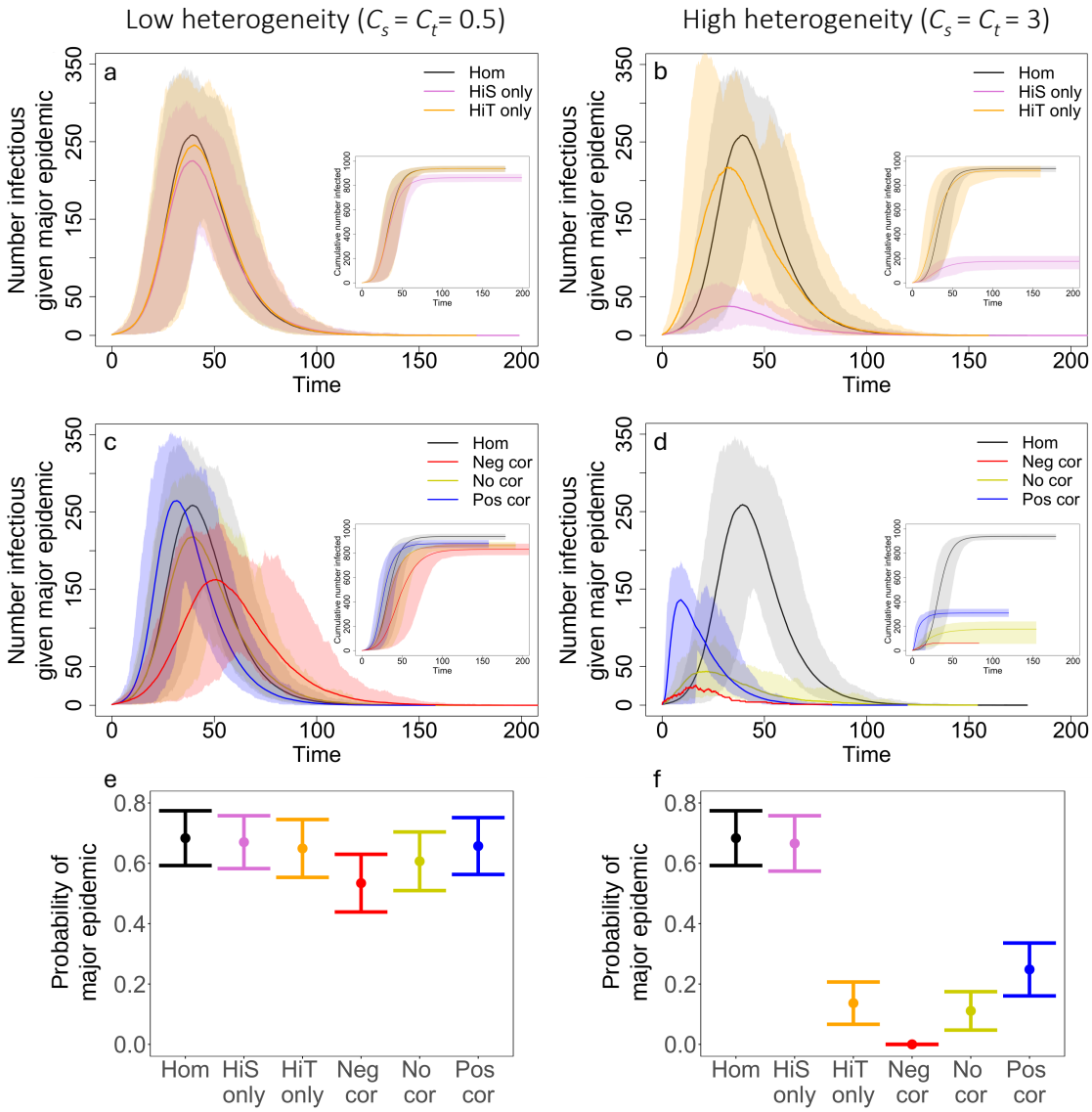

Figure 5: Correlations affect the disease dynamics in our model. The plots (a-d) show the average number of infected individuals over the course of an epidemic. Shaded regions represent 95% CIs determined from the major epidemics of 200 simulations for homogeneity (black), heterogeneity in susceptibility only (pink), heterogeneity in transmission only (orange), negative correlation (red, $\rho = -1$), no correlation (yellow, $\rho = 0$), and positive correlation (blue, $\rho = 1$). (e,f) The average probability of a major epidemic for each case with 200 simulations and error bars of +/- 2 standard deviations. (Insets) The cumulative number of individuals infected with 95% CIs from the major epidemics of 200 simulations. $C_s = C_t = 0.5$ in (a,c,e), $C_s = C_t = 3$ in (b,d,f), $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

had smaller peak sizes and smaller final epidemic sizes than epidemics without heterogeneity in susceptibility (i.e., homogeneity; Figs 9 and 10). Also as expected, we found that heterogeneity in transmission alone ($C_t = 0.5, 1$, or 3 and $C_s = 0$) reduced the probability of a major epidemic (Fig 7). When epidemics did occur, they had the same peak and final epidemic sizes as the homogeneity case (Figs 9 and 10).

In a population heterogeneous with regard to both transmission and susceptibility, we found that when transmissibility and susceptibility are uncorrelated, the effect on disease dynamics is the same as combining the effects of each individual source of heterogeneity alone. In particular, when uncorrelated, the epidemics had a similar peak size and final epidemic size to those with heterogeneity in susceptibility alone (respectively, Figs 9 and 10). Additionally, they had a similar probability of a major epidemic as with heterogeneity in transmission alone (Fig 7).

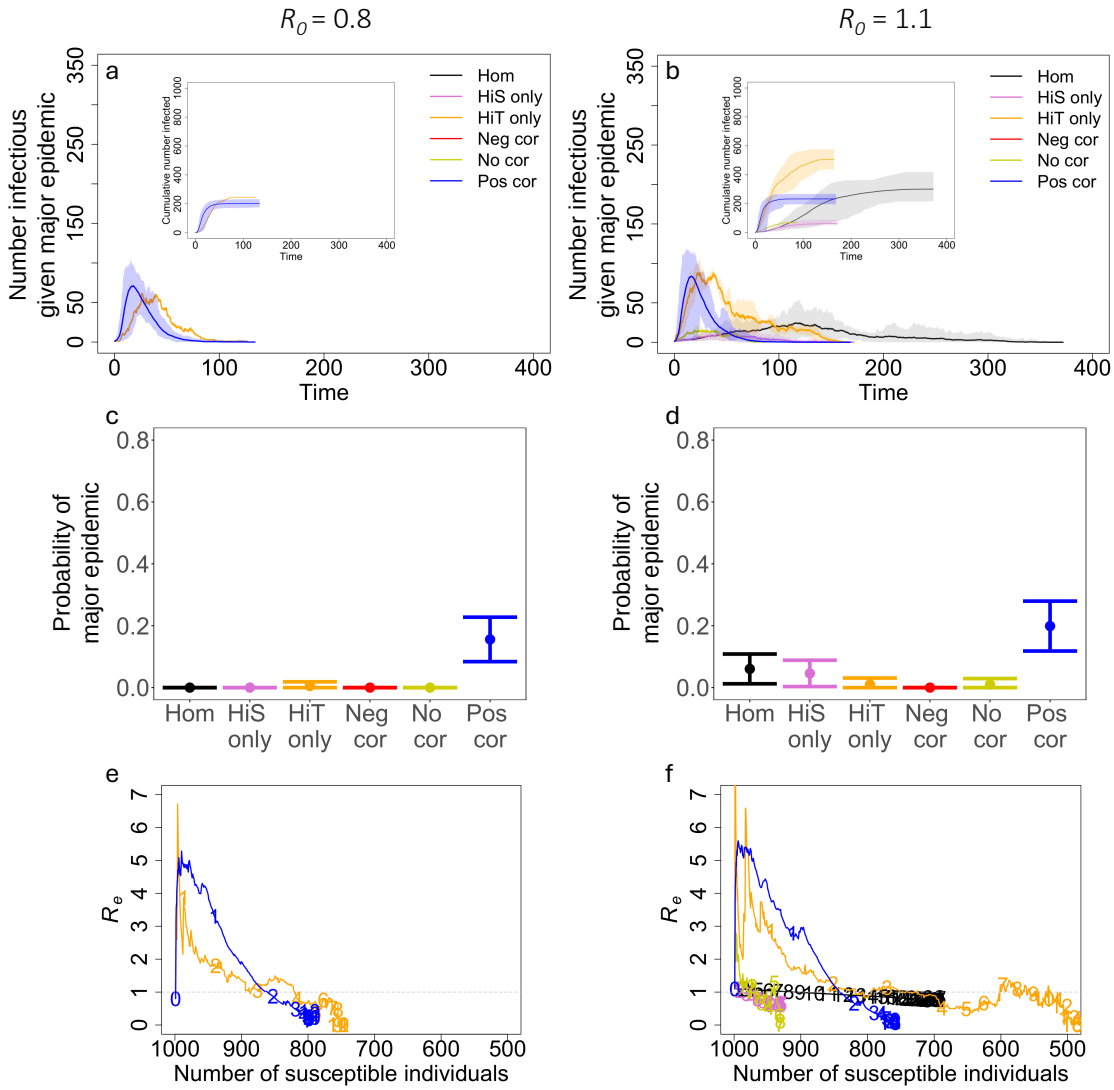

Figure 6: Positive correlations can lead to larger, more likely major epidemics when R_0 is small. (a,b) The average number of infected individuals over the course of an epidemic. Shaded regions represent 95% CIs determined from the major epidemics of 200 simulations for homogeneity (black), heterogeneity in susceptibility only (pink), heterogeneity in transmission only (orange), negative correlation (red, $\rho = -1$), no correlation (yellow, $\rho = 0$), and positive correlation (blue, $\rho = 1$). (c,d) The average probability of a major epidemic for each case with 200 simulations and error bars of 2xSD. (e,f) R_e plotted against the number of susceptible individuals (S) averaged over the major epidemics from 200 simulations. The numbers on each trajectory represent time in the epidemic for every 10 units of time starting from the left (e.g., 1 is placed at time $t = 10$, 2 at $t = 20$, etc.). The dotted gray line shows $R_e = 1$. (Insets) The cumulative number of individuals infected with 95% CIs from the major epidemics of 200 simulations. $C_s = C_t = 3$, $R_0 = 0.8$ in (a,c,e), $R_0 = 1.1$ in (b,d,f), $N = 1000$, and $I_0 = 1$.

On the other hand, when there are correlations between transmissibility and susceptibility, these heterogeneities have interactive effects on disease dynamics. Since positive correlations indicate that the most susceptible individuals are also the most transmissible, when the first infected individual is transmissible enough to spread infection, epidemics readily take off and rapidly spread (Fig 5c and 5d). In this situation, however, epidemics also burn out quickly because the more susceptible individuals become infected early leaving only individuals that are lowly susceptible and lowly transmissible. In comparison to the uncorrelated case and the case without heterogeneity, a positive correlation results in a fast epidemic that quickly peaks and crashes. A positive correlation makes the overall probability of an epidemic more likely than in the uncorrelated case. Compared to the case where there is no heterogeneity, a positive correlation can make an epidemic less likely (Fig 5e and 5f) or more likely (Fig 6c and 6d) depending on the value of R_0.

Since negative correlations indicate that the most susceptible individuals tend to be the least transmissible, and vice versa, this type of correlation between susceptibility and transmissibility causes epidemics to be hard to start and hard to sustain. This pattern arises because the individuals most likely to become infected are those least likely to cause new infections. When these epidemics do take off, in comparison to both the uncorrelated case and the case without heterogeneity, they are smaller, and they tend to either die out quickly or slowly meander through the population depending on the transmissibility of individuals infected early in the epidemic (Fig 5c and 5d).

As expected from previous literature (Lloyd-Smith et al., 2005) and our results with heterogeneity in transmission alone, the probability of a major epidemic is greatly influenced by the absolute magnitude of the heterogeneity in transmission (Fig 7). When holding correlations constant, increasing heterogeneity in transmission greatly decreases the chance of a major epidemic, but this effect disappears for subcritical R_0 (Fig 8). Notably, correlations between susceptibility and transmissibility can partially offset the effect of heterogeneity in transmission, as demonstrated in Fig 7. In that figure and in Fig 5e and 5f, major epidemics occur more often with high heterogeneity in transmission and a positive correlation than with the same level of heterogeneity in transmission and no correlation (for example, the case where $C_t = 3$, $C_s = 3$, and $\rho = 0$ vs. $\rho = 1$). With decreased R_0, the effect of correlations is more pronounced. Positive correlations frequently lead to a greater chance of a major epidemic than in both the absence of correlation and the absence of heterogeneity with $R_0 = 1.1$ (Supplement Fig S2). Moreover, with subcritical $R_0 = 0.8$ and a positive correlation, major epidemics could still occur up to 25% of the time under the parameter sets explored in Fig 8. In the absence of correlations, there is little to no impact on the probability of a major epidemic from the level of heterogeneity in susceptibility.

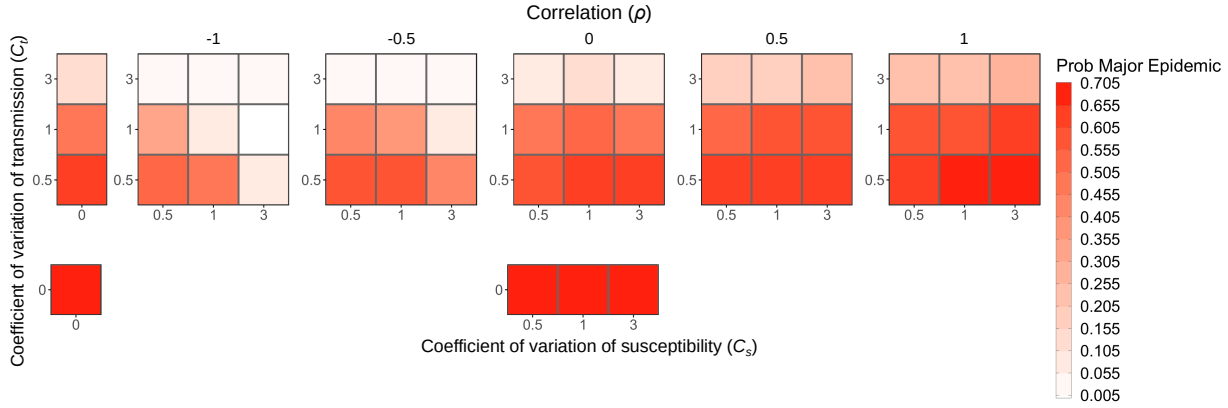

Figure 7: The probability of a major epidemic decreases as the level of heterogeneity in transmission increases, increases with positive correlation, and decreases with negative correlation. Each box is shaded to show the probability of a major epidemic, as defined in the text, averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. While heterogeneity in transmission primarily determines the probability of a major epidemic, there are also effects from correlation and heterogeneity in susceptibility. Positive correlation results in a higher probability, negative correlation results in a lower probability, and increased levels of heterogeneity in susceptibility further decrease the chance of a major epidemic when there is a negative correlation. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

Also, as expected from previous literature (Aguas et al., 2020; Gomes et al., 2022; Montalbán et al., 2022) and our results with heterogeneity in susceptibility alone, in the absence of correlations, the peak size and final epidemic size is reduced as the level of heterogeneity in susceptibility increases, while these features are largely independent of heterogeneity in transmission (Figs 9 and 10). When susceptibility and transmissibility are correlated, however, changing the magnitude of heterogeneity in transmission can have large impacts on peak size and final epidemic size. In particular, with negative correlations, increasing heterogeneity in transmission decreases peak size and final epidemic size and decreases it relative to the uncorrelated case. With positive correlations, on the other hand, increasing heterogeneity in transmission increases peak size in comparison to the uncorrelated case. This is because at high level of heterogeneity in transmission, the

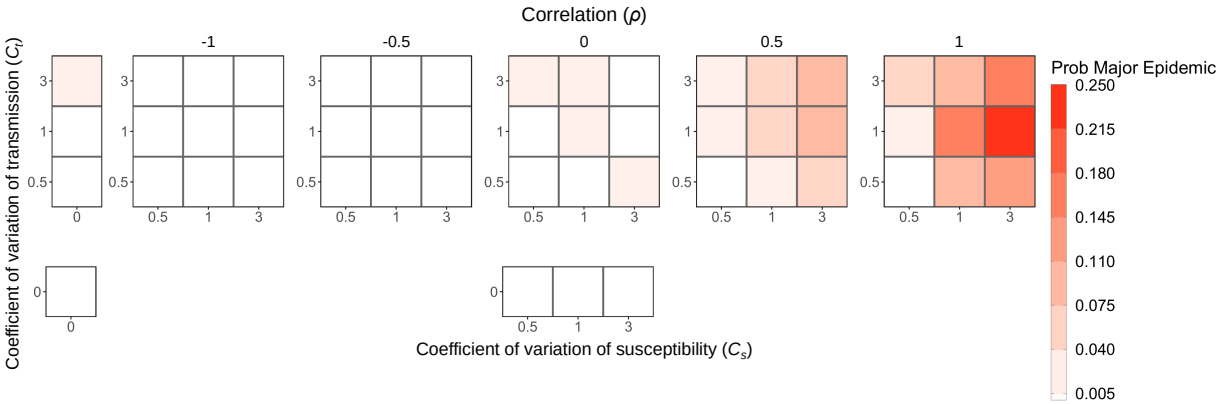

Figure 8: Major epidemics are still consistently possible with subcritical R_0 when there is positive correlation. Each box is shaded to show the probability of a major epidemic, as defined in the text, averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. $N = 1000$, $I_0 = 1$, and $R_0 = 0.8$.

individuals most likely to become infected early tend to be those that are most transmissible. Moreover, with positive correlations, it is often the case that peak sizes are larger than in the case where there is no heterogeneity in susceptibility. In such cases, the final epidemic size tends to remain smaller, but the reverse is true for small R_0 (Supplement Figs S4 and S12). These two features arise with a positive correlation because the epidemic sweeps through the population quickly and crashes, typically resulting in a large peak but ultimately smaller epidemic (except when R_0 is small). Negative correlations on the other hand lead to smaller peaks and final epidemic sizes than both the no correlation and homogeneous cases. This is because epidemics are hard to sustain with negative correlations since the individuals that are more likely to become infected tend to be less likely to transmit onward.

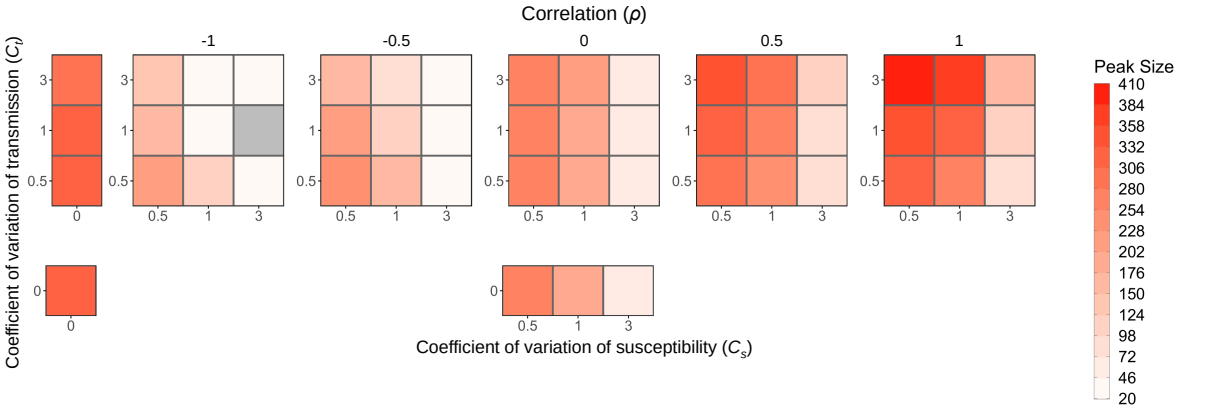

Figure 9: Peak size decreases as the level of heterogeneity in susceptibility increases, increases with positive correlation, and decreases with negative correlation. Each box is shaded to show the peak size averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so the average in each box is based on a different sample size. The gray box represents a parameter combination that resulted in no major epidemics as defined in the text. While heterogeneity in susceptibility primarily determines the peak size, there are also effects from correlation and heterogeneity in transmission. Positive correlation results in a larger peak size, negative correlation results in a smaller peak size, and increased levels of heterogeneity in transmission enhance the effect of correlation. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

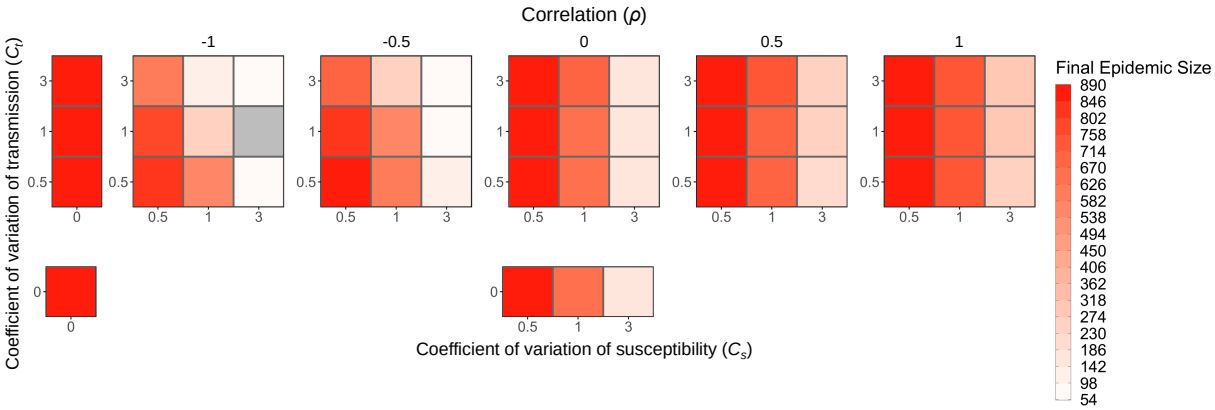

Figure 10: Final epidemic size decreases as the level of heterogeneity in susceptibility increases, increases with positive correlation, and decreases with negative correlation. Each box is shaded to show the final epidemic size averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so the average in each box is based on a different sample size. The gray box represents a parameter combination that resulted in no major epidemics as defined in the text. While heterogeneity in susceptibility primarily determines the final epidemic size, there are also effects from correlation and heterogeneity in transmission. Positive correlation results in a larger epidemic size, negative correlation results in a smaller epidemic size, and increased levels of heterogeneity in transmission enhance the effect of correlation. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

In contrast to the other epidemic features, which were driven primarily by levels of heterogeneity in either susceptibility or transmission, the peak time of an epidemic is strongly modified by the direction and magnitude of correlates between susceptibility and transmissibility rather than by their level of heterogeneity (Fig 11). With positive correlations, epidemics take off quickly, so the peak tends to occur earlier in comparison to both the no correlation case and no heterogeneity case. With negative correlations, epidemics that take off tend to do so slowly, so the peak tends to occur later than in the no correlation and no heterogeneity cases. Epidemics that successfully take off in this negative correlation case, however, often start with a highly transmissible individual, so the peak can be earlier or later depending on the level of heterogeneity in transmission.

Another way to show the impact of correlation on the timing of an epidemic is to explore the time to the jth infection. Taking this approach reveals that the direction and magnitude of correlations between susceptibility and transmissibility can drive the timing of the jth infection, with patterns generally similar to those of peak time (Fig 12). In Fig 13, we simplify the data from Fig 12 to show only the time to the 50th infection, which allows us to more readily compare across correlations and levels of heterogeneity. We also looked at the time for the epidemic to reach 50% of its final size (i.e., $j = 0.5(N - S_{\text{final}})$) to compare the speed of epidemics with different sizes and found generally similar results (Supplement Fig S1). Compared to the case with heterogeneity but no correlation, a positive correlation leads to the 50th infection happening earlier and a negative correlation leads to the 50th infection happening later. Notably, a positive correlation also leads to a faster epidemic than in the homogeneous case, but as can be seen in Fig 12, this changes for large values of j (i.e., where the jth infection occurs close to the end of the epidemic). As shown in Fig 12, a negative correlation results in a slower epidemic than other cases except when there is high heterogeneity in transmission ($C_t = 3$). In this case, the time to the jth infection for negative correlation is earlier than both the no correlation and homogeneity cases for small j but is later for large j. This is because the first infected individual must have high transmissibility for a major epidemic to occur, but the epidemic slows down over time. Fig 12 additionally shows that the trajectories for heterogeneity in transmission alone and no correlation tend to match those for homogeneity and heterogeneity in susceptibility alone respectively, which occurs because the difference in timing is primarily driven by heterogeneity in susceptibility or correlations between susceptibility and transmissibility.

In addition to these measures, we looked at the effect of correlations on the effective reproductive number

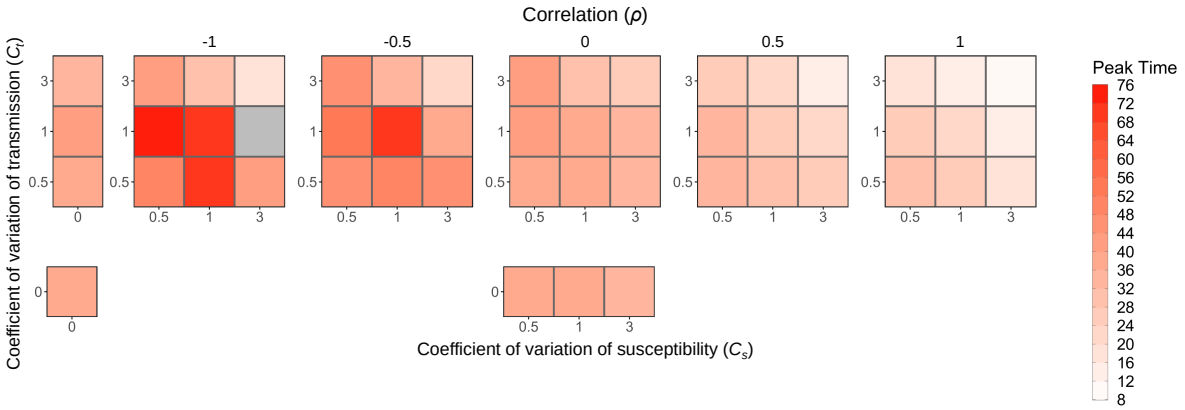

Figure 11: Peak time is earlier with positive correlation and later with negative correlation. Each box is shaded to show the peak time averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so the average in each box is based on a different sample size. The gray box represents a parameter combination that resulted in no major epidemics as defined in the text. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

R_e over the course of a major epidemic, by plotting R_e versus the number of susceptible individuals in the population S (Fig 14). This plot additionally shows how R_e changes over time. In the case with no heterogeneity, R_e is linearly related to the size of the susceptible population. When there is heterogeneity in transmission only, the average R_e follows the same trajectory with respect to S as in the homogeneous case although there is variation between outbreak simulations. When there is heterogeneity in susceptibility only, R_e declines faster with respect to S than in the case with no heterogeneity, but it follows the same trajectory as the no correlation case. This is because the presence of heterogeneity in susceptibility creates an infection selection process whereby highly susceptible individuals are infected early in the epidemic, resulting in an increasingly less susceptible population and therefore smaller R_e over time. The dynamics are fundamentally altered by positive and negative correlations between susceptibility and transmissibility. In the presence of a positive correlation, R_e shoots up at the start of an epidemic before rapidly declining faster than linearly against S. In addition, the trajectory plays out faster in time than in the other cases examined. This is because the epidemic quickly sweeps through the most susceptible and transmissible individuals before rapidly dying out. In the presence of a negative correlation, R_e immediately drops then rapidly declines as a function of S. The exception to this is that when there is high heterogeneity in transmission ($C_t = 3$), R_e first shoots up before dropping. This happens because major epidemics only occur in this case when the first infected individual has high transmissibility. The rate of decline of R_e as a function of time, however, is slower with negative correlations than in the other cases after the initial precipitous drop. These patterns arise because individuals that are more susceptible do not contribute as much to onward transmission and vice versa, resulting in a slower epidemic. With $C_t = 3$, however, major epidemics may be faster for negative correlations because the individual infected at the start must have high transmissibility for the epidemic to take off.

5. Discussion

An important question in epidemiology is why epidemics follow different trajectories. It is well known that heterogeneities in transmission and susceptibility can individually impact disease dynamics in known ways (Dwyer et al., 1997; Lloyd-Smith et al., 2005; Gomes et al., 2014; Langwig et al., 2017; Gomes et al., 2022). Here we have shown that when both heterogeneities are present simultaneously, correlations between transmissibility and susceptibility can drastically alter disease dynamics. We showed this by first conducting a systematic review to identify what was previously known and what knowledge gaps remained, and second by creating a stochastic, individual-based SIR model to generate results to fill in the knowledge gaps. Our literature review found that the effects of correlations between transmissibility and susceptibility were

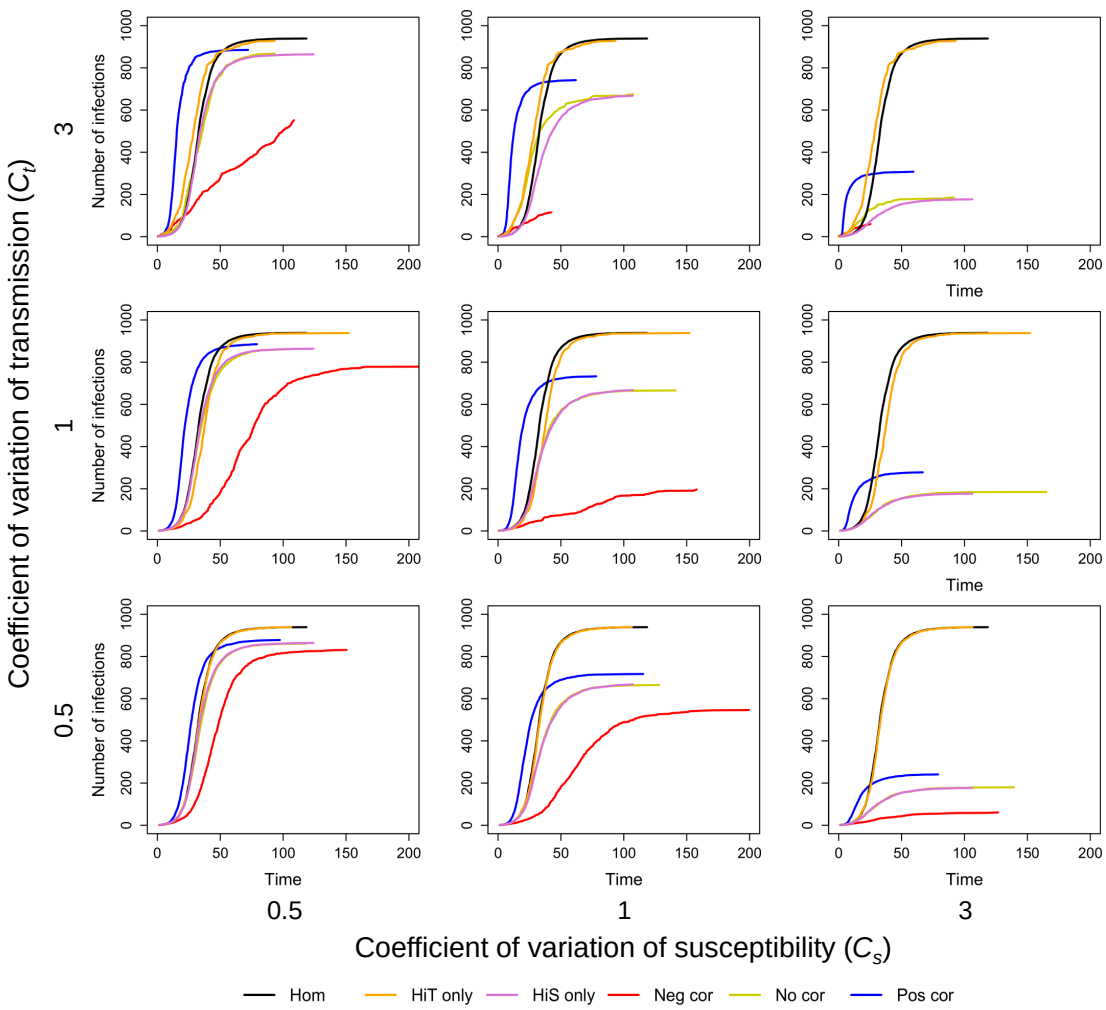

Figure 12: Time to the jth infection is earlier with positive correlation and later with negative correlation. The plots show the median time to the jth infection (where j is the value on the y-axis) from the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. Each plot includes trajectories for the cases of homogeneity (black), heterogeneity in transmission alone (orange), heterogeneity in susceptibility alone (purple), perfect negative correlation ($\rho = -1$, red), no correlation (yellow), and perfect positive correlation ($\rho = 1$, blue). The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so each line is based on a different sample size. There is no line for negative correlation with $C_t = 1$ and $C_s = 3$ because this parameter combination resulted in no major epidemics as defined in the text. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

understudied, particularly with regard to how the dynamics of models with these correlations differed from those of models with heterogeneities but no correlations. Using our model, we found that if the drivers of heterogeneity in transmission and susceptibility are independent, then the effects on disease dynamics are essentially the same as each heterogeneity source alone (i.e., heterogeneity in transmission reduces the probability of a major epidemic, and heterogeneity in susceptibility reduces the peak and final epidemic sizes). In contrast, the impacts on disease dynamics are more complicated when susceptibility and transmissibility are correlated. Positive correlations result in major epidemics that are larger, faster, and more likely in comparison to the case with no correlation (Fig 15). In comparison to the case with no heterogeneity, epidemics are faster with positive correlations. When R_0 is a bigger value, positive correlations lead to smaller and less likely major epidemics than the dynamics with homogeneity (Fig 5), but when R_0 is close to or less than 1, positive correlations can lead to larger and more likely major epidemics (Fig 6). Negative correlations result in major epidemics that are less likely and smaller than both the no correlation and homogeneous cases, but these epidemics can be faster or slower than both cases depending on the level of heterogeneity in

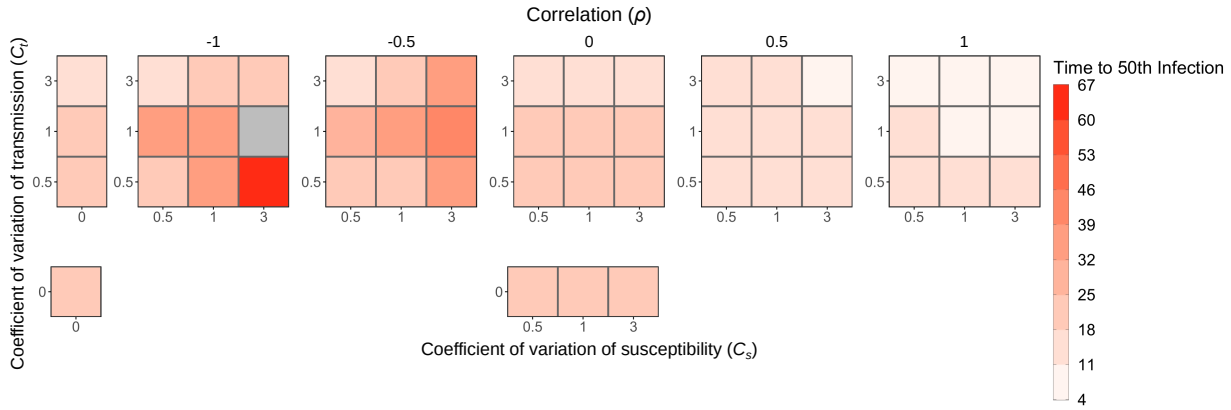

Figure 13: Time to the 50th infection is earlier with positive correlation and later with negative correlation. Each box is shaded to show the time to the 50th infection averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so the average in each box is based on a different sample size. The gray box represents a parameter combination that resulted in no major epidemics as defined in the text. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

transmission (Fig 12).

From our systematic literature review, we have identified that few published studies explored the impact of correlation between transmissibility and susceptibility on disease dynamics (Cardell and Kanouse, 1989; Becker and Marschner, 1990; Daley et al., 2000; Andreasen, 2011; Clancy and Pearce, 2013; Hickson and Roberts, 2014; Tkachenko et al., 2021; Kawagoe et al., 2021; Allard et al., 2023). The studies that do exist tend to have compared models with heterogeneity and correlations to models without heterogeneity, meaning that they were largely unable to assess the impact of the correlation itself on disease dynamics. Moreover, the studies that have been published used different model structures, parameter values, and assumptions, resulting in disagreements on the effects of correlations not just in magnitude but also in direction. We found that the majority of models agreed that positive correlations resulted in smaller final epidemic sizes and earlier peaks in comparison to cases without heterogeneity (Cardell and Kanouse, 1989; Daley et al., 2000; Andreasen, 2011; Clancy and Pearce, 2013; Hickson and Roberts, 2014; Tkachenko et al., 2021; Kawagoe et al., 2021). In two models cases, however, this did not occur (Cardell and Kanouse, 1989; Hickson and Roberts, 2014). For one of the models, Cardell and Kanouse (1989) kept the growth rate constant for the first month of the epidemic to compare between cases, resulting in a later peak as individuals not infected during the first month are more likely to be both less susceptible and less transmissible. In the other case, Hickson and Roberts (2014) kept the mean product of transmissibility and susceptibility constant across cases and thus found positive correlations resulted in a smaller epidemic size but a similar peak time compared to homogeneity. Our literature review also found disagreement between the effects of negative correlations on disease dynamics. For instance, in comparison to models without heterogeneity, some models found that negative correlations could result in a smaller (Daley et al., 2000; Andreasen, 2011; Clancy and Pearce, 2013; Hickson and Roberts, 2014), larger (Andreasen, 2011; Clancy and Pearce, 2013; Hickson and Roberts, 2014), or similar (Andreasen, 2011; Clancy and Pearce, 2013) final epidemic size. This may be because of the respective authors' choices of distributions used to model susceptibility and transmissibility or the metrics kept constant to make comparisons across cases. Lastly, our literature review revealed that the vast majority of prior studies only compared models with correlations to models without heterogeneity, and thus neglected the comparison to models with heterogeneity but no correlation between transmissibility and susceptibility. The two that compared the case with correlations to the case without correlations restricted their analysis to only final epidemic size and a positive correlation, meaning that there were substantial knowledge gaps in understanding how correlation between susceptibility and transmissibility itself impacted disease dynamics (Fig 4).

There has recently been renewed interest in understanding how heterogeneity in and interactions between

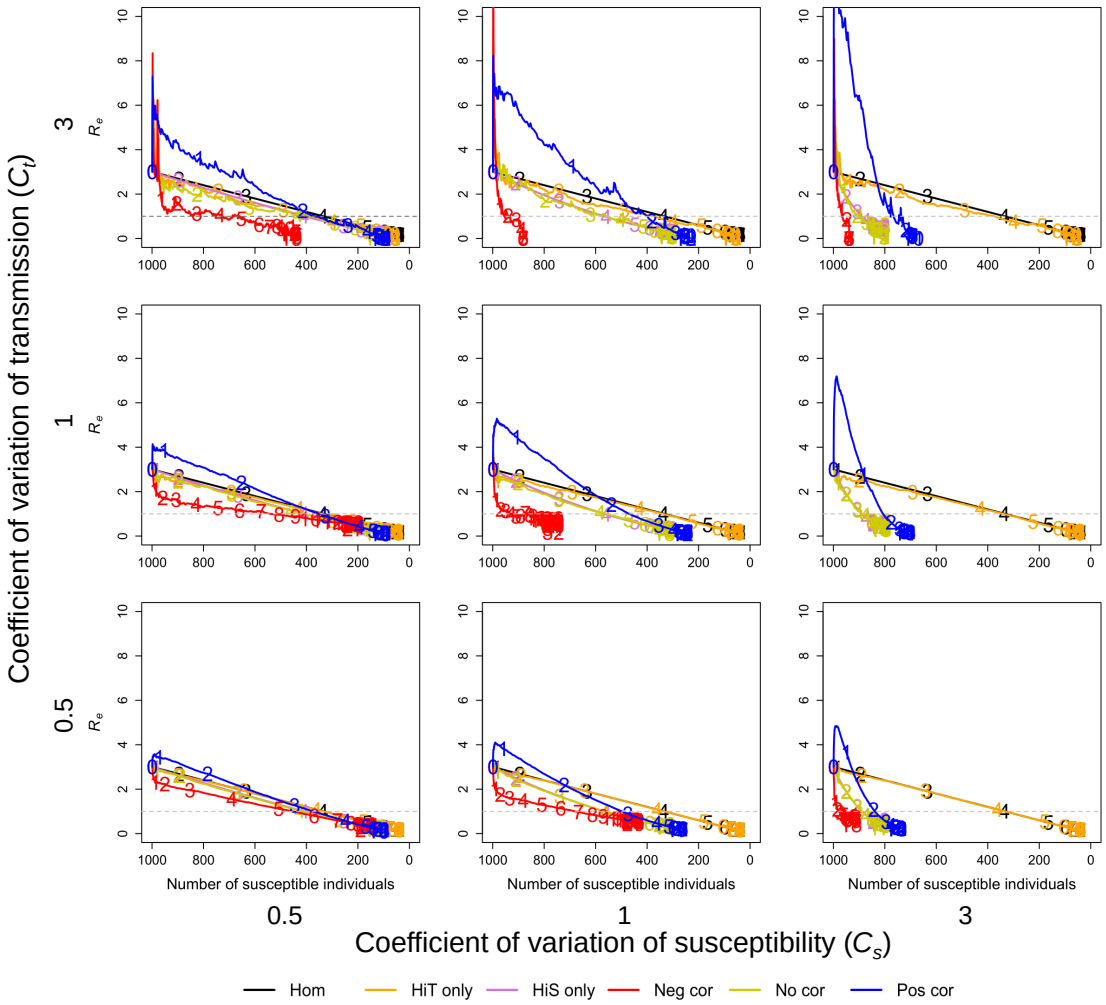

Figure 14: The effective reproductive number R_e depends on both the levels of heterogeneity and the correlation. The plots show R_e plotted against the number of susceptible individuals (S) averaged over the major epidemics from 200 simulations for various levels of heterogeneity in transmission (C_t), heterogeneity in susceptibility (C_s), and the correlation between transmissibility and susceptibility. Each plot includes trajectories for the cases of homogeneity (black), heterogeneity in transmission alone (orange), heterogeneity in susceptibility alone (purple), perfect negative correlation (red, $\rho = -1$), no correlation (yellow, $\rho = 0$), and perfect positive correlation (blue, $\rho = 1$). The numbers on each trajectory represent time in the epidemic for every 10 units of time starting from the left (e.g., 1 is placed at time $t = 10$, 2 at $t = 20$, etc.). The dotted gray lines show $R_e = 1$. The number of simulations that were major epidemics, which can be determined by the probability of a major epidemic in Fig 7, was different for each parameter combination, so each line is based on a different sample size. There is no line for negative correlation with $C_t = 1$ and $C_s = 3$ because this parameter combination resulted in no major epidemics as defined in the text. $N = 1000$, $I_0 = 1$, and $R_0 = 3$.

traits related to transmissibility and susceptibility affect disease dynamics. Beagle et al. (2022) investigated how correlations between traits related to transmissibility and superspreading affect peak epidemic size. They found that, in comparison to the case without correlations, positive correlations between contact rate and infectiousness result in larger peak sizes, whereas negative correlations result in smaller peak sizes. We believe that this result arose in their analysis because, despite their focus on transmissibility, contact rate also implicitly impacts susceptibility. Their results thus further support our finding that correlations between susceptibility and transmissibility impact peak size (increase with positive correlations and decrease with negative correlations).

In a recent preprint, Harris et al. (2024) also investigated the effects of correlations between transmissibility and susceptibility using truncated Gaussian distributions with correlations ranging from -0.6 to 0.6. Largely in agreement with our results, they found that in comparison to the case with no correlation between

susceptibility and transmissibility, positive correlations lead to faster, stronger, more likely epidemics, whereas negative correlations lead to slower, weaker, less likely epidemics. Our analysis differs from theirs in that we used gamma distributions for our heterogeneity, we analyzed the impact of correlations ranging from -1 to 1, and we investigated a wider range of values for the level of heterogeneity. In addition, we explored the effect of altering R_0, which demonstrated that positive correlations can lead to larger major epidemics compared to the homogeneous case even in cases where no outbreak would occur in the absence of heterogeneity.

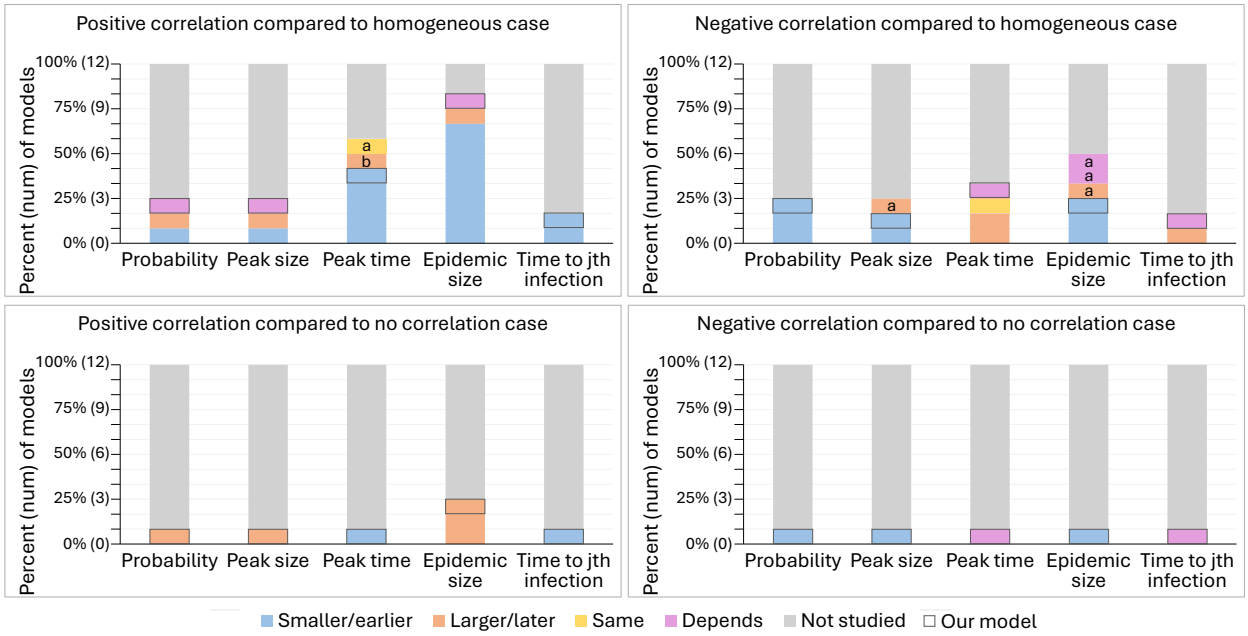

Figure 15: Results from our model fill in knowledge gaps about the effects of correlation from the systematic review. These plots show the effects of positive and negative correlations between transmissibility and susceptibility on the probability of a major epidemic, peak size, peak time, final epidemic size, and time to the jth infection in comparison to disease dynamics with homogeneity or no correlation according to 11 models from the 9 studies included in the systematic review and our model. The effect of correlation is classified for each model and measure as resulting in an attribute that is smaller/earlier (blue), larger/later (orange), the same (yellow), any of these results depending on the levels of heterogeneity or value of R_0 (purple), or not studied (gray) in comparison to the homogeneous or no correlation case. The results from our model have a black outline. For positive correlation compared to the homogeneous case, we found that the effect of correlation on the probability of a major epidemic and the final epidemic size depends on R_0 (defined by Eq 4), and the effect of correlation on the peak size depends on R_0 and the level of heterogeneity. When R_0 is close to or less than 1 ($R_0 = 0.8$ or 1.1), positive correlation results in a larger probability, peak, and final epidemic size than with homogeneity, whereas when R_0 is increased ($R_0 = 3$), positive correlations results in a less likely and smaller epidemic than with homogeneity. With $R_0 = 3$, the peak size can still be larger for positive correlations than homogeneity when there is high heterogeneity in transmission and low heterogeneity in susceptibility. Also, note that the time to the jth infection is earlier for positive correlations than homogeneity under the condition that j is not large (i.e., j is not close to the final epidemic size). For negative correlation, we found that the effect of correlation on both the peak time and the time to the jth infection depends on the level of heterogeneity in transmission. High heterogeneity leads to an earlier peak and jth infection compared to the homogeneous and no correlation cases while low heterogeneity leads to these attributes being later. The letters 'a' and 'b' denote results from the systematic review models that are inconsistent with our results where 'a' means the model associated with that result defined R_0 according to Eq. 5 and 'b' means the model kept a constant growth rate for the first month.

R_0, or the basic reproductive number, is a benchmark parameter used to evaluate the threat posed by an emerging infectious disease. While this number can be, and has been, valuable for predicting and managing epidemics (Ferguson et al., 2001; Lipsitch et al., 2003; Hens et al., 2015), its interpretation can change in the presence of heterogeneity. The probability of a major epidemic can be calculated based on R_0, but this probability must be adjusted if populations have heterogeneity in transmission (Lloyd-Smith et al., 2005). Likewise, the final epidemic size and threshold for herd immunity can be calculated based on R_0, but these values must be adjusted to account for heterogeneity in susceptibility (Ball, 1985; Montalbán et al., 2022). We have shown that when susceptibility and transmissibility are correlated, further adjustments are needed.

These issues are exacerbated by small details that impact the calculation of R_0. As we discussed above, R_0 can be defined as the product of mean transmissibility and mean susceptibility when individuals are equally likely to be the first infected (Eq 4), and R_0 can also be alternatively defined by the mean product of transmissibility and susceptibility when the likelihood of being the first infected is exactly proportional to an individual's susceptibility (Eq 5). These values need not be similar and holding one versus the other constant can lead to fundamentally different conclusions regarding the effects of correlations between susceptibility and transmissibility on disease dynamics. Four of the 11 models in our systematic review made comparisons across disease dynamics by keeping the initial mean transmissibility and mean susceptibility constant (i.e., R_0 defined as in Eq 4). Five of the 11 models made comparisons by keeping the initial mean product of transmissibility and susceptibility constant (i.e., R_0 defined as in Eq 5). This reflects the ambiguity present in defining and interpreting R_0. For our analysis, we used the first definition of R_0, electing to keep the mean force of infection and mean risk of infection constant between different levels of correlations and heterogeneity. One way this played out is that the effective reproductive number R_e immediately differed between the positive, negative and uncorrelated cases. It also led to the counterintuitive results that major epidemics could still occur, and with relatively high probability, when $R_0 < 1$ in the case of positive correlations because R_e typically became greater than 1 after the first transmission event. Hence, R_0, as defined by Eq. 4, is a poor measure of outbreak potential when there are correlations between transmissibility and susceptibility.

We propose that incorporating heterogeneity and correlations into disease models may improve predictive capability, particularly in cases where there are known or likely correlations between transmissibility and susceptibility. For instance, variation in the number of contacts that individuals have is a frequently assumed mechanism for generating heterogeneity in transmission (Lloyd-Smith et al., 2005). In such cases, there is an implicit positive correlation between susceptibility and transmissibility since contact rate impacts both exposure and onward transmission. However, other sources of heterogeneity can be difficult to estimate, especially early in an epidemic. This is slowly changing, since recent increased interest in real-time disease forecasting has resulted in the development of novel methods to estimate heterogeneity in susceptibility (Gomes et al., 2022; Anderson et al., 2023; Tuschhoff and Kennedy, 2024). We propose that it will likewise be important to develop methods to measure correlations between susceptibility and transmissibility.

Based on our finding that positive correlations lead to larger, faster, more likely major epidemics, we suggest that increased surveillance for disease outbreaks may be warranted in places or subpopulations where positive correlations are likely. Likewise, public health interventions can be targeted towards settings where positive correlations are likely, possibly even leading to negative correlations. An example of this would be targeting limited vaccines towards individuals with many contacts. This strategy has been proposed previously for the purpose of reducing superspreading events without consideration for correlation (Vidondo et al., 2012; Manzo and van de Rijt, 2020; Saunders and Schwartz, 2021). If such correlations were present, taking these correlations into account would increase the support in favor of utilizing this approach.

One consideration of our study is that we investigated the effect of correlation between two variables (the force of infection and risk of infection) on disease dynamics. In reality, there may be many interacting variables with varying degrees and directions of correlation. For example, there may be correlations between factors associated with infection duration, such as recovery rate or virulence, and those important for disease transmission, such as the force of infection and risk of infection that we investigated. The impact of this added complexity may affect disease dynamics differently than what we found (Beagle et al., 2022) and should be explored further.

Overall, the correlation between transmissibility and susceptibility is an important consideration for epidemiological models since all of our epidemic measures were affected by correlation to some extent. Although the probability of a major epidemic, peak size, and final epidemic size were greatly affected by the levels of heterogeneity, positive correlations generally counteracted the effects of heterogeneity and negative correlation generally enhanced them. The impact of a positive correlation is especially important for populations with small or subcritical R_0 where it is still possible to have sizable major epidemics even when no such epidemic would occur without a correlation or in the absence of heterogeneity. The timing and speed of an epidemic (i.e., the peak time and time to the jth infection) were driven largely by correlations, with a positive correlation leading to a faster epidemic and a negative correlation leading to a slower one. The levels of heterogeneity also impacted timing, particularly early in the epidemic, where increased heterogeneity in transmission led to a more explosive epidemic, and increased heterogeneity in susceptibility led to a faster epidemic as infection swept through the more susceptible individuals. As all of these measures of disease

dynamics are wrapped up with R_e, both correlation and the levels of heterogeneity strongly influenced R_e.

The 2022 mpox epidemic may provide a real-world example of the effect of correlations. In that outbreak, the number of mpox cases rapidly increased and then rapidly declined after only a small fraction of the overall population was infected (Murayama et al., 2024). Mpox is spread mainly through close contact, resulting in a perfect positive correlation between susceptibility and transmissibility (Laurenson-Schafer et al., 2023), assuming no change in behavior following infection. While other factors may have contributed, such as information campaigns and targeted interventions (Delaney et al., 2022), the rapid rise and subsequent sharp decline of mpox is consistent with the dynamics caused by positive correlations between transmissibility and susceptibility. It remains an open question to determine the degree to which the control of mpox was due to public health measures versus the epidemic running its course.

Acknowledgements: We thank Ephraim Hanks for feedback on our model setup.

References

Aguas, R., Corder, R.M., King, J.G., Goncalves, G., Ferreira, M.U., Gomes, M.G.M., 2020. Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. medRxiv .

Allard, A., Moore, C., Scarpino, S.V., Althouse, B.M., Hébert-Dufresne, L., 2023. The role of directionality, heterogeneity, and correlations in epidemic risk and spread. SIAM Review 65, 471–492.

Anderson, R.M., May, R.M., 1991. Infectious diseases of humans: dynamics and control. Cambridge Univer-sity Press.

Anderson, T.L., Nande, A., Merenstein, C., Raynor, B., Oommen, A., Kelly, B.J., Levy, M.Z., Hill, A.L., 2023. Quantifying individual-level heterogeneity in infectiousness and susceptibility through household studies. Epidemics 44. 100710.

Andreasen, V., 2011. The final size of an epidemic and its relation to the basic reproduction number. Bulletin of Mathematical Biology 73, 2305–2321.

Baker, M.G., Peckham, T.K., Seixas, N.S., 2020. Estimating the burden of United States workers exposed to infection or disease: a key factor in containing risk of COVID-19 infection. PloS One 15, e0232452.

Ball, F., 1985. Deterministic and stochastic epidemics with several kinds of susceptibles. Advances in Applied Probability 17, 1–22.

Beagle, A.S., Budischak, S.A., Craft, M.E., Forbes, K., Hall, R., Nguyen, D., Cressler, C.E., 2022. Do some super-spreaders spread better? effects of individual heterogeneity in epidemiological traits. medRxiv , 2022–04.

Becker, N., Marschner, I., 1990. The effect of heterogeneity on the spread of disease, in: Stochastic Processes in Epidemic Theory: Proceedings of a Conference held in Luminy, France, October 23–29, 1988, Springer. pp. 90–103.

Bielicki, J.A., Duval, X., Gobat, N., Goossens, H., Koopmans, M., Tacconelli, E., van der Werf, S., 2020. Monitoring approaches for health-care workers during the COVID-19 pandemic. The Lancet Infectious Diseases 20, e261–e267.

Cardell, N.S., Kanouse, D.E., 1989. Modeling heterogeneity in susceptibility and infectivity for HIV infection, in: Mathematical and Statistical Approaches to AIDS Epidemiology. Springer, pp. 138–156.

Clancy, D., Pearce, C.J., 2013. The effect of population heterogeneities upon spread of infection. Journal of Mathematical Biology 67, 963–987.

Collins, J.P., Campbell, A.P., Openo, K., Farley, M.M., Cummings, C.N., Hill, M., Schaffner, W., Lindegren, M.L., Thomas, A., Billing, L., et al., 2020. Outcomes of immunocompromised adults hospitalized with laboratory-confirmed influenza in the united states, 2011–2015. Clinical Infectious Diseases 70, 2121–2130.

Daley, D., Gani, J., Yakowitz, S., 2000. An epidemic with individual infectivities and susceptibilities. Mathematical and Computer Modelling 32, 155–167.

Delaney, K.P., Sanchez, T., Hannah, M., Edwards, O.W., Carpino, T., Agnew-Brune, C., Renfro, K., Kachur, R., Carnes, N., DiNenno, E.A., Lansky, A., Ethier, K., Sullivan, P., Baral, S., Oster, A.M., 2022. Strategies adopted by gay, bisexual, and other men who have sex with men to prevent monkeypox virus transmission—united states, august 2022. MMWR. Morbidity and mortality weekly report 71.

Dwyer, G., Elkinton, J.S., Buonaccorsi, J.P., 1997. Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. The American Naturalist 150, 685–707.

Ericksen, K.P., Trocki, K.F., 1992. Behavioral risk factors for sexually transmitted disease in American households. Social Science & Medicine 34, 843–853.

Ferguson, N.M., Donnelly, C.A., Anderson, R.M., 2001. Transmission intensity and impact of control policies on the foot and mouth epidemic in great britain. Nature 413, 542–548.

Galvin, S.R., Cohen, M.S., 2004. The role of sexually transmitted diseases in HIV transmission. Nature Reviews Microbiology 2, 33–42.

Gomes, M.G.M., Ferreira, M.U., Corder, R.M., King, J.G., Souto-Maior, C., Penha-Gonçalves, C., Gonçalves, G., Chikina, M., Pegden, W., Aguas, R., 2022. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. Journal of Theoretical Biology 540. 111063.

Gomes, M.G.M., Lipsitch, M., Wargo, A.R., Kurath, G., Rebelo, C., Medley, G.F., Coutinho, A., 2014. A missing dimension in measures of vaccination impacts. PLoS Pathogens 10, e1003849.

Harris, J.D., Gallmeier, E., Dushoff, J., Beckett, S.J., Weitz, J.S., 2024. Infections are not alike: the effects of covariation between individual susceptibility and transmissibility on epidemic dynamics. medRxiv , 2024–10.

Heffernan, J.M., Smith, R.J., Wahl, L.M., 2005. Perspectives on the basic reproductive ratio. Journal of the Royal Society Interface 2, 281–293.

Hens, N., Abrams, S., Santermans, E., Theeten, H., Goeyvaerts, N., Lernout, T., Leuridan, E., Van Kerckhove, K., Goossens, H., Van Damme, P., et al., 2015. Assessing the risk of measles resurgence in a highly vaccinated population: Belgium anno 2013. Eurosurveillance 20.

Hickson, R., Roberts, M., 2014. How population heterogeneity in susceptibility and infectivity influences epidemic dynamics. Journal of Theoretical Biology 350, 70–80.

Kalajdzievska, D., Li, M.Y., 2011. Modeling the effects of carriers on transmission dynamics of infectious diseases. Mathematical Biosciences & Engineering 8, 711–722.

Karvonen, A., Jokela, J., Laine, A.L., 2019. Importance of sequence and timing in parasite coinfections. Trends in Parasitology 35, 109–118.

Kawagoe, K., Rychnovsky, M., Chang, S., Huber, G., Li, L., Miller, J., Pnini, R., Veytsman, B., Yllanes, D., 2021. Epidemic dynamics in inhomogeneous populations and the role of superspreaders. Physical Review Research 3, 033283.

Langwig, K.E., Wargo, A.R., Jones, D.R., Viss, J.R., Rutan, B.J., Egan, N.A., Sá-Guimarães, P., Kim, M.S., Kurath, G., Gomes, M.G.M., Lipsitch, M., 2017. Vaccine effects on heterogeneity in susceptibility and implications for population health management. mBio 8, e00796–17.

Laurenson-Schafer, H., Sklenovská, N., Hoxha, A., Kerr, S.M., Ndumbi, P., Fitzner, J., Almiron, M., de Sousa, L.A., Briand, S., Cenciarelli, O., et al., 2023. Description of the first global outbreak of mpox: an analysis of global surveillance data. The Lancet Global Health 11, e1012–e1023.

Li, L., Hsu, S.H., Wang, C., Li, B., Sun, L., Shi, J., Ren, Y., Wang, J., Zhang, X., Liu, J., 2021. Characteristics of viral pneumonia in non-hiv immunocompromised and immunocompetent patients: a retrospective cohort study. BMC Infectious Diseases 21, 1–12.

Lipsitch, M., Cohen, T., Cooper, B., Robins, J.M., Ma, S., James, L., Gopalakrishna, G., Chew, S.K., Tan, C.C., Samore, M.H., et al., 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science 300, 1966–1970.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., Getz, W.M., 2005. Superspreading and the effect of individual variation on disease emergence. Nature 438, 355–359.

Manzo, G., van de Rijt, A., 2020. Halting SARS-CoV-2 by targeting high-contact individuals. arXiv .

Medzhitov, R., Schneider, D.S., Soares, M.P., 2012. Disease tolerance as a defense strategy. Science 335, 936–941.

Memoli, M.J., Athota, R., Reed, S., Czajkowski, L., Bristol, T., Proudfoot, K., Hagey, R., Voell, J., Fiorentino, C., Ademposi, A., et al., 2014. The natural history of influenza infection in the severely immunocompromised vs nonimmunocompromised hosts. Clinical Infectious Diseases 58, 214–224.

Montalbán, A., Corder, R.M., Gomes, M.G.M., 2022. Herd immunity under individual variation and reinfection. Journal of Mathematical Biology 85. 2.

Murayama, H., Pearson, C.A., Abbott, S., Miura, F., Jung, S.m., Fearon, E., Funk, S., Endo, A., 2024. Accumulation of immunity in heavy-tailed sexual contact networks shapes mpox outbreak sizes. The Journal of Infectious Diseases 229, 59–63.

Nunner, H., van de Rijt, A., Buskens, V., 2022. Prioritizing high-contact occupations raises effectiveness of vaccination campaigns. Scientific Reports 12, 737.

Park, S.Y., Kim, T.H., Lee, E., Loeb, M., Jeong, Y.S., Kim, J.H., Oh, S.M., Cheong, S., Park, H., Jo, S., et al., 2022. A SARS-CoV-2 outbreak associated with vaccine breakthrough in an acute care hospital. American Journal of Infection Control 50, 1006–1012.

Permar, S.R., Griffin, D.E., Letvin, N.L., 2006. Immune containment and consequences of measles virus infection in healthy and immunocompromised individuals. Clinical and Vaccine Immunology 13, 437–443.

Saunders, H.A., Schwartz, J.M., 2021. COVID-19 vaccination strategies depend on the underlying network of social interactions. Scientific Reports 11, 24051.

Shaw, C.L., Kennedy, D.A., 2021. What the reproductive number R_0 can and cannot tell us about COVID-19 dynamics. Theoretical Population Biology 137, 2–9.

Southall, E., Ogi-Gittins, Z., Kaye, A., Hart, W., Lovell-Read, F., Thompson, R., 2023. A practical guide to mathematical methods for estimating infectious disease outbreak risks. Journal of Theoretical Biology 562, 111417.

Tkachenko, A.V., Maslov, S., Elbanna, A., Wong, G.N., Weiner, Z.J., Goldenfeld, N., 2021. Time-dependent heterogeneity leads to transient suppression of the COVID-19 epidemic, not herd immunity. Proceedings of the National Academy of Sciences 118, e2015972118.

Tuschhoff, B.M., Kennedy, D.A., 2024. Detecting and quantifying heterogeneity in susceptibility using contact tracing data. PLOS Computational Biology 20, e1012310.

Vidondo, B., Schwehm, M., Bühlmann, A., Eichner, M., 2012. Finding and removing highly connected individuals using suboptimal vaccines. BMC Infectious Diseases 12, 1–8.

Zélé, F., Magalhães, S., Kéfi, S., Duncan, A.B., 2018. Ecology and evolution of facilitation among symbionts. Nature Communications 9, 4869.