

Mitigating the spread of SARS-CoV-2 in prisons and other congregate settings

Seth Blumberg, Ada Kwan, Chris Hoover, Phoebe Lu, David Sears, ...others from CALProt

5/1/21

1 Abstract

Because SARS-CoV-2 has led to explosive outbreaks in congregate settings such as prisons, the need for effective outbreak prevention and mitigation strategies for these settings is critical. While many transmission models have been developed for community spread, less attention has been given to modeling the interdependence of disease introduction and spread seen in congregate settings. Here we consider how interventions that decrease the size of the susceptible populations, such as vaccination and/or depopulation, impact the expected number of infections. Introduction of disease into the resident population from the community is modeled as a branching process, while spread between residents is modeled via a compartmental model. Control is modeled as a proportional decrease in both the number of susceptible residents and the reproduction number. Using the California state prison system as a case example, we find that vaccination and/or decarceration can have a greater than linear effect on anticipated infections (e.g. For a reproduction of 3.0, reducing the size of the susceptible population by 20% reduced overall disease burden by 47%). We explore implications for the optimal distribution for housing residents in two different buildings, and when a reactive vaccination approach is preferable to a preemptive one when vaccine supply is limited. These findings provide a quantitative framework that may help policy-makers optimize the implementation of infection control procedures in congregate settings.

2 Introduction

Given the transmissibility and morbidity associated with SARS-CoV-2, there is a pressing need to identify strategies for disease control. A particular challenge is congregate settings such as prisons, nursing homes, and crowded workplaces where transmission is amplified **citations needed**. The increased risk of transmission results in a higher potential for an outbreak where a large number of people can become infected within a few weeks. Besides the higher caseload in these settings, the cases in these settings often have a higher prevalence of

co-morbidities that contribute to worse disease outcomes [Citation needed](#). Thus an outbreak can result in a sudden surge of hospital admissions that can strain healthcare capacity and seed increased transmission within the wider community [1, 2]. Outbreaks in prison settings are further complicated by the additional security, training, and contractual resources needed to hospitalize an incarcerated person.

One opportunity to reduce the public health risk associated with congregate settings, is decreasing the number of susceptible individuals via vaccination and/or depopulation [2, 3]. [See comments in overleaf](#) Reducing the number of susceptible individuals can both decrease the chance of an outbreak occurring and the size of any outbreaks that occur [4]. [Maybe give an example of how depopulation has been achieved - e.g. decarceration through the expedited release of prisoners?](#) A second opportunity to reduce disease spread is by optimizing the allocation of [residents](#) amongst multiple facilities. [Are there any examples we can use showing how these decisions have been made](#) A third opportunity to optimize disease control is to evaluate the benefit of preemptive versus reactive control efforts. For example when vaccine supply is limited, it may sometimes be preferable to vaccinate as many residents as possible preemptively, and other times it may be preferable to implement a rapid vaccination campaign in facilities that are experiencing an outbreak.[greenAda - see overleaf comment](#)

These three opportunities for preventing and mitigating COVID-19 outbreaks are the focus of these manuscript. Quantifying their impact and optimizing their implementation are vital as there has been incredible spread of infection in congregate settings despite [masking](#), restricting movement of residents, and screening staff at entry. To provide a quantitative framework to evaluate these public health interventions, we describe a model for the probability of an outbreak occurring in a congregate setting within a specified time period, as well as the size of an outbreak that may occur. To illustrate the implications of our model, we focus on publicly available data from the California Department of Corrections and Rehabilitation (CDCR).

3 Methods

3.1 Data

COVID-19 data for all state prison facilities operated by the CDCR are reported daily in a CDCR data dashboard. Machine readable time series of these daily reports were acquired from the UCLA COVID Behind Bars project which gathers and organizes COVID-19 data from carceral facilities across the United States, including CDCR. [5] [Cite: data downloaded 2021-02-02 https://github.com/uclalawcovid19behindbars/data](#) Time series of incident cases were derived by taking the daily difference of reported cumulative cases and removing anomalies such as negative case counts. [Chris : Can you explain data cleaning more. Does ignoring negative case numbers than contribute to extra cases in the next data point?](#)

3.2 Model overview

We break up outbreak dynamics in congregate settings into three stages (Figure 1). First, a case has to be introduced into the congregate setting. This may occur due to direct transfer of an infected case into the congregate population, or from transmission from a worker or visitor. Second, an introduction of a case (or a few cases) can either be self-limited or progress to a full outbreak. The probability of an outbreak occurring is impacted by the stochastic nature of disease spread, the reproduction number, and the degree of transmission heterogeneity. The reproduction number, R_0 , is the average number of cases each new case causes when all of their contacts are susceptible to disease. We use the dispersion parameter, k , to denote the amount of transmission heterogeneity. Third, if an outbreak is established, the overall impact is in proportion the total number of cases infected. Since the large number of cases overwhelms the stochasticity of transmission, this third stage is more deterministic in nature.

We assume that $R_0 > 1$, because when R_0 is less than one, transmission is self-limited and outbreaks that involve a large portion of the population are not expected. The key outputs for each stage of the model is the rate of introductions, ϕ , the probability that an introduction results in an outbreak, P_{ob} and the expected size of any outbreaks that occur, I_{tot} . Our models for each stage of an outbreak are described in the supplementary methods.

We choose a population size for the residents of **1,000 as this is on the order of the average size of the CDCR prison facilities**. The prevalence of disease in the staff, the number of staff-resident contacts and the probability that an infected staff transmits disease to a resident during a contact were estimated to 0.01%, 10 and 1% respectively based on a combination of community dynamics and the empirical observation of relatively frequent outbreaks occurring in the CDCR facilities [red]References needed. We may need to update estimates. Based on current literature, we assume an incubation period of three days(i.e. time between the occurrence of infection and the onset of infectiousness), and an infectious period of seven days.[6–10] **see four refs in commented text**

Since our results focus on relative proportions of residents affected in different control scenarios, the exact values for the aforementioned parameters will not impact our results. A notable exception is that the value of the reproduction number is of particular importance. We thus explore a range of values for the reproduction number as there is great variability in the literature. Population-level estimates of the reproduction number are as low as 1.5, but there are also published estimates of the reproduction number as high as eight in prison settingsRef. For the purpose of our analyses, R_0 is defined within the context of system-wide control interventions that are in place before consideration of vaccination and/or depopulation. That is R_0 may incorporate interventions such as masking, and limiting movements that are expected to occur regardless of additional interventions designed to reduce the number of susceptible residents.

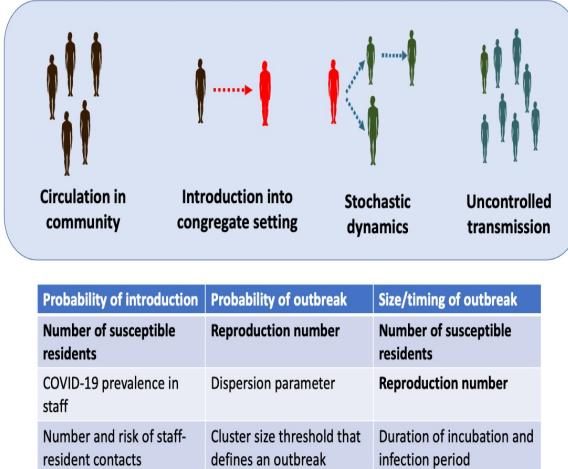


Figure 1: **Stages of an outbreak.** For an outbreak to occur in a congregate setting, an infection that circulates in the community must be introduced into the congregate setting. Then stochastic dynamics determine whether or not an outbreak occurs. Once an outbreak occurs, uncontrolled transmission dictates the size and time course of the outbreak. The table lists the variables that are used in each step of the model. The bolded variables are impacted by a reduction in the number of susceptible individuals via vaccination and/or depopulation.

3.3 Impact of preemptive vaccination and/or depopulation

Each of the stages of our model for outbreak dynamics is influenced by either the number of susceptible individuals, R_0 , or both. This serves as the basis for assessing how the reduction of the susceptible population by vaccination and/or depopulation can impact the burden of disease in congregate settings.

We define N_0 , and R_0 as the size of the congregate population and the reproduction number prior to any control interventions that decrease the number of susceptibles, and prior to any outbreak occurring. We let γ represent the amount of ‘control’ as determined by the proportional reduction in the size of the susceptible population that occurs with vaccination and/or depopulation,

$$N = (1 - \gamma) N_0. \quad (1)$$

We assume a linear relationship between the effective reproduction number R , and γ ,

$$R = (1 - \Theta\gamma) R_0. \quad (2)$$

Here we have defined a susceptibility index, Θ . For $\Theta = 0$, the reproduction number remains constant even when the number of susceptible residents are reduced. This would be equivalent to a vaccine that eliminates disease but has no impact on transmission. In contrast when $\Theta = 1$, the ratio between the size of the susceptible population and the reproduction number remains constant. This corresponds to a vaccine that is equally effective at reducing disease and transmission in a well-mixed population. We assume that the dispersion parameter is independent of γ .

To model the overall impact of depopulation, we define E to be the average number of cases expected due to outbreaks that are initiated over a defined time interval, T . The expected number of introductions is ϕT . The overall probability, P_{ob} of an outbreak occurring is one minus the probability that all introductions lead to extinctions. The average number of cases expected due to outbreak is the overall probability that an outbreak occurs times the expected size of an outbreak. That is,

$$P_{ob} = 1 - p_e^{\phi T} \quad (3)$$

$$E = P_{ob} \cdot I_{tot} \quad (4)$$

Here we have assumed that only one outbreak can occur in a given residence because enough individuals would be infected so that there would be subsequent herd immunity. Each of ϕ , P_{ob} , and I_{tot} depend on γ . The impact of control can be probed by evaluating how E depends on γ .

3.4 Optimizing distribution of residents

When considering the optimal proportion of residents to house in each of two buildings, we specify the baseline occupancy of each building, N_0^1 and N_0^2 , as well as the basic reproduction numbers at those baseline occupancies, R_0^1 and R_0^2 . We also specify the overall vaccination rate, V . For a proportion σ of residents assigned to building one, we set the size of the susceptible pool and the effective reproduction numbers prior to any outbreaks as,

$$N_S^1 = \sigma \cdot (N_0^1 + N_0^2) \quad (5)$$

$$N_S^2 = (1 - \sigma) \cdot (N_0^1 + N_0^2) \quad (6)$$

$$R_S^1 = R_0^1 \cdot \frac{N_S^1 \cdot (1 - V)}{N_0^1} \quad (7)$$

$$R_S^2 = R_0^2 \cdot \frac{N_S^2 \cdot (1 - V)}{N_0^2} \quad (8)$$

Except for these parameterizations, the model is run for each building independently to determine the building-specific outbreak probabilities, P_{ob}^1 and P_{ob}^2 , and outbreak sizes, I_{tot}^1 and I_{tot}^2 . The joint probability of an outbreak occurring in at least one building and the overall expected number of cases is then,

$$P_{ob}^{all} = P_{ob}^1 + P_{ob}^2 - P_{ob}^1 \cdot P_{ob}^2 \quad (9)$$

$$I_{tot}^{all} = P_{ob}^1 \cdot I_{tot}^1 + P_{ob}^2 \cdot I_{tot}^2 \quad (10)$$

3.5 Reactive vaccination and/or depopulation

As an alternate to preemptive control, we consider reactive approaches. In this scenario, we allow the baseline model to run until a ‘trigger threshold’ is reached for the number of cases. The model continues to run unchanged for a time period specified by the vaccine delay. Then, the susceptible population experiences a one time relative decrease of one minus the vaccine efficacy. The rest of the model proceeds as before.

4 Results

4.1 Evaluating model assumptions

From the beginning of California Department of Corrections and Rehabilitation reporting in April of 2020 through April 30 2021 there were a total of ~~xxx~~ SARS-CoV2 cases in 34 prison facilities. Within each facility, periods of low disease prevalence were interrupted by focal outbreaks caused by rapid spread (Figure 2). Some facilities have multiple outbreaks, which might occur if separate residential buildings have outbreaks at different times. Importantly, the periods of low disease prevalence can have small clusters of cases that do not progress to an outbreak. Overall this data supports our key assumptions that ~~there~~ outbreak dynamics are supported by sporadic introductions into the residential community and that stochastic dynamics determine the probability that a sporadic introductions progresses into a large outbreak.

4.2 Stages of outbreak dynamics

A prerequisite for an outbreak to occur is introduction of disease into the residential community. For our model, the frequency of introductions increases with a larger residential population, increased prevalence of disease in the community, a higher resident-staff contact rate, or higher probability that a resident contact with an infected staff causes disease (Figure S6).

A disease introduction may or may not lead to an outbreak. The probability of an outbreak increases as the number of introductions, or the reproduction number increases (Figure S7). High values of the dispersion parameter correspond to homogeneous transmission and more predictable dynamics, whereas low values correspond to heterogeneous dynamics that are more likely to produce either explosive outbreaks or dead-ends to transmission. Thus high values of the dispersion parameter lead to a higher outbreak probability (Figure S7). Numerous studies have shown that respiratory diseases tend to exhibit superspreading behavior and recent studies indicate that SARS-CoV-2 follow this pattern. [11–14]. Cite LloydSmith 2005, Something on SARS-1 and MERS superspreading, althouse-stochasticity-2020, adam2020clustering, susswein-characterizing-2020. Based on the estimates of the dispersion parameter these studies have found for

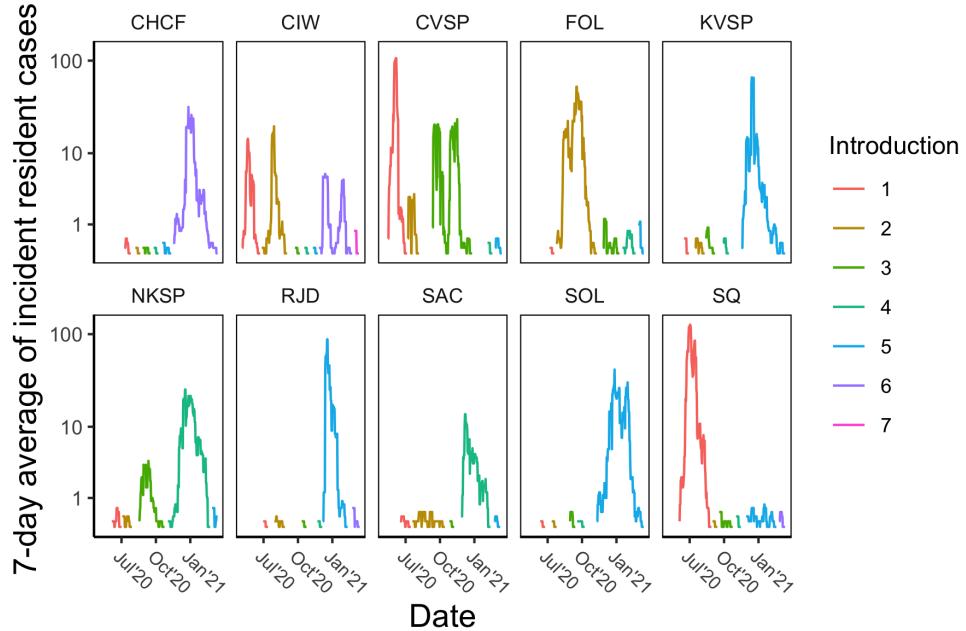


Figure 2: COVID-19 incidence in California state prison facilities. Each panel represents one of the 34 state facilities. For visualization purposes, Y axes are log transformed and 7-day rolling averages of incident counts are displayed. To highlight the stochastic impact of disease introductions, a new color is used for the ~~prevalence~~ data whenever there is a period of no cases lasting at least ~~10~~ days. Thus the different colors approximate the consequence of individual disease introductions into the residential community. The panels represent facilities where at least five introductions ~~are modeled to have~~ occurred. **Chris:** Per hidden comments in overleaf, consider using 14-day washout instead

SARS-CoV-2 and other respiratory diseases, we henceforth assume a value of 0.2 for the dispersion parameter.

Once an outbreak occurs, the number of infectious individuals in our model grows exponentially until there is a significant depletion of susceptible individuals (Figure S9). A reproduction number above one is necessary for an outbreak to occur, but even values moderately above one lead to a large ~~proportion of infected residents~~. A reproduction number as low as 1.5 results in over half of the residents being infected, and a reproduction number of 2.5 leads to a large ~~majority of cases~~ being infected (as seen by the asymptotic value for the number of removed individuals in the two panels of Figure S9).

As the number of susceptible residents are reduced by depopulation or vaccination, the consequences of an outbreak are mitigated (Figure S9, upper left panel). Particularly for values of R above 3, the ~~proportion of residents who are~~

sick remains high even if 50% of the susceptible population is removed (Figure S9, upper right panel). Thus for higher values of R , most of the decrease in the size of the outbreak is driven by the decrease in the number of **residents that can get sick** rather than by a reduction in the transmissibility of virus. Although not a prime focus of our analysis, reducing the size of the susceptible population also delays the timing of the peak for the number of infectious individuals and reduces the maximum number of individuals infected at once (Figure S9, bottom panels). This can help facilitate roll-out of public health interventions that can further decrease the burden of disease and minimize overcrowding of healthcare facilities treating severe cases.

4.3 Impact of preemptive vaccination and/or depopulation

When our models of disease introduction, outbreak probability and outbreak size are combined, there is a significant impact of interventions that decrease the size of the susceptible population (Figure 3). Given that there is likely substantial variability in the value of the reproduction number with estimates as high as eight, it is notable that control interventions have a large impact for a wide array of R_0 values. Ref: prison R of 8 For example, when the susceptible population is decreased by 20% and transmission decreases linearly with control (i.e. $\Theta = 1$), the expected number of total infections decreases by 73%, 47%, 40%, and 38%, for R_0 of 1.5, 3.0, 5.0, and 8.0 respectively.

In general, the expected number of infections decrease with **higher levels of control**. As **control** increases, the absolute number of expected infections decreases faster with higher levels of R_0 because the **baseline numbers are higher**. However, the relative change in expected infections decrease faster for lower R_0 . Notably, the reduction in the number of cases occurs in greater than linear fashion (seen by colored lines falling below the black line in Figure 3). The decrease is even more significant when control impacts both the susceptibility to infection and the transmission potential of individuals, as compared to control affecting susceptibility alone (seen by solid lines falling below the dashed lines in Figure 3). The greater than linear impact of control can be explained by how the expected number of infections is the product of the probability of an introduction, the probability that an introduction leads to an outbreak and the size of an outbreak. All of these factors depends on either the effective reproduction number and/or the size of the susceptible population. Since control via vaccination and/or depopulation decreases the size of the susceptible population and possibly the effective reproduction number as well, each factor contributes to the reduction in the number of cases when control is applied (Figures 1 and S10). When the dependency of these individual components of outbreak dynamics are multiplied together, the result is a great than linear impact of control.

4.4 Optimizing distribution of residents

I wonder if this subsection gets too technical?

Besides, decreasing the size of the susceptible population, the probability of an outbreak and the expected number of cases from an outbreak can also be reduced by optimizing the allocation of residents between buildings (Figures 4 and S11). The optimal proportion that is ideal for two sets of buildings depends on the relative transmissibility in each building, and the overall number of susceptible residents.

To minimize the probability that an outbreak occurs, the optimal distribution equalizes the reproduction number between the buildings. Thus if the buildings have the same transmission potential, the outbreak probability is minimized by keeping the building occupancy balanced at 50% each (seen by x-value of solid dots in **scenarios 1,3,4, and 6** of Figure 4). However in Scenarios 2 and 5, the outbreak probability is minimized when 25% of residents are in building 1 and 75% are in building 2 because that makes the reproduction number equal in both buildings (e.g. R is 3 and 2.4 for scenarios 2 and 5 respectively).

The optimal proportion of residents to house in each building to minimize the expected number of cases from outbreaks does not follow a simple rule. For scenarios with moderate values for the probability of an outbreak occurring, the optimal proportion to house in each building to minimize the expected number of cases is closer to 50% than what is expected from analyzing the outbreak probability alone (scenarios 2 and 5 in Figure 4). For scenarios with low values of the probability of an outbreak occurring, the expected number of cases due to outbreaks occurring can remain low for a wide range of the proportion housed in each building (Scenario 4 in 4). This is because the effective reproduction number can be kept below one across this range and thus outbreaks are not expected to occur. For high risk scenarios where the probability of an outbreak occurring is close to one, the best opportunity to minimize the number of cases is to reduce the number of residents in one building to the point that a large outbreak can only occur in the more populous building (seen by the optimal distribution being just 19% in building 2 for scenario 6 in the bottom panel of Figure 4). However, the benefit of the optimal distribution in the high risk scenario is quite marginal.

4.5 Preemptive versus reactive control

When resources are constrained, there may be a tradeoff between preemptive and reactive strategies for outbreak control. For example, suppose there are five equally sized facilities and only enough vaccine for 20% of the total population. As a preemptive approach, 20% of the residents in each facility could be vaccinated. An alternative reactive approach would be to wait until an **outbreak is detected** in a building and then vaccinate all residents in the building as quickly as possible. Visualization of the outcomes for these strategies shows that the effectiveness of a reactive strategy depends heavily on how quickly vaccination can occur and how effective it is (Figure 5). A reactive vaccine is more likely to be superior to a preemptive one if the reproduction number is low. The higher the reproduction number is, the quicker a reactive strategy will need to be deployed in order for it to have much impact. For example, even with a long

vaccine delay our model predicts a reactive vaccine efficacy of 20% would be superior to preemptive vaccination rate of 10% for an R_0 of 1.5 (seen by the red line being below the dotted line in the left panel of Figure 5). However, if R_0 were 3.5, even a rapid reactive vaccination with an efficacy of 20% would not yield better results than preemptive vaccination of 10% (seen by the red line being above the dotted line in the right panel of Figure 5).

A reactive depopulation strategy is also possible, provided that the residents who are removed from the outbreak are quarantined prior to moving to a new location. Otherwise depopulation risk seeding outbreaks in new locations.

5 Discussion

The burden of COVID-19 within prisons, nursing homes and other congregate settings has been devastating (Figure 2). Although the prevalence of COVID-19 is decreasing in many areas, the risk remains high in hotspots throughout the world. In addition, new variants of SARS-CoV-2 could lead to new outbreaks, and it is unclear whether the population immunity that currently exists will persist in the months and years to come. Thus, to prevent unnecessary death, disease, and economic loss policies are needed to prevent outbreaks from occurring and to mitigate outbreaks that have already started.

Our results provide perspective on several opportunities for disease control in congregate settings. First, our results highlight the utility of reducing the size of the susceptible population via depopulation and vaccination (Figure 10). Small changes in the in the susceptible population can have greater impact on the overall number of cases due to the combination of decreasing the rate of disease introductions, the probability that an introduction leads to an outbreak and the size of any outbreak that occurs (Figure 1)

Consideration of the spatial and temporal nature of outbreak dynamics provides further opportunities for disease control. For example, we discovered that the disease burden can typically be minimized by dividing residents into buildings in a way that equalizes the reproduction number in each building (Figure 4). However, if the outbreak potential is very high, then it can be extremely difficult to control an outbreak once it begins. Thus for situations in rich transmission risk is particularly high, it may be prudent to create small protected zones for residents who have a high likelihood of developing severe infection. Meanwhile, since outbreaks evolve over time, there may be opportunities to mitigate the final size of an existing outbreak by rapidly depopulating or vaccinating residents in the affected population (Figure 5) The success of this reactive mode of control will depend on the speed with which implementation can be achieved, and thus pre-pandemic planning is imperative.

In order to frame our analysis of control strategies in a manner that was intuitive and transparent, we made many assumptions about the transmission dynamics. Depending on specific circumstances of spread in congregate settings, some assumptions will likely be more relevant than others. Specifically we ignore the possibility of direct transfer of infected individuals from one building to

another as can inadvertently happen for individuals who are transferred during the incubation period **Citation?**. We ignore the fluctuation of disease prevalence in the staff, particularly as surges may occur due to staff-staff transmission or resident-to-staff transmission. Similarly we ignore how staff-resident contacts may evolve over the course of a pandemic, particularly as access to PPE and education about infection control improves over time. Our use of a SEIR transmission model assumes that within a building everyone is in equal contact with each other and thus ignores the finer scale structure of population dynamics within a building. Our SEIR model also incorporates the standard assumptions of compartmental models including constant exponential rates of transitions and standard mass-action transmission. Finally, we have assumed a linear relationship between the size of the susceptible population and the reproduction number, but factors such as the relationship between duration-of-exposure and infection status may yield a non-linear relationship.

blueAda: I am not sure where to include your insightful comments about distinguishing between population-level and individual-level control, as well as listing the various control options (see commented text in overleaf). I was thinking this might be a good spot?

6 Conclusion

Congregate settings pose a risk of large disease outbreaks. To reduce the burden that outbreaks have on residents, staff and the community at large, it is important to optimize strategies for preventing and mitigating outbreaks. We find that preemptive reduction of the size of the susceptible population via depopulation, vaccination and/or other prophylactic measures can have a greater than linear effect on outcomes. Additional strategies that can be effective include optimizing the location of resident to reduce the overall probability and expected size of an outbreak. Reactive strategies such as rapid vaccination of all residents once an outbreak has begun can also be beneficial, provided the reproduction number is not too high and the control interventions are implemented quickly.

References

- [1] Eric Reinhart and Daniel L. Chen. “Incarceration And Its Disseminations: COVID-19 Pandemic Lessons From Chicago’s Cook County Jail”. eng. In: *Health Affairs (Project Hope)* 39.8 (Aug. 2020), pp. 1412–1418. ISSN: 1544-5208. DOI: 10.1377/hlthaff.2020.00652.
- [2] Andre G. Montoya-Barthelemy et al. “COVID-19 and the Correctional Environment: The American Prison as a Focal Point for Public Health”. eng. In: *American Journal of Preventive Medicine* 58.6 (June 2020), pp. 888–891. ISSN: 1873-2607. DOI: 10.1016/j.amepre.2020.04.001.

- [3] Brandy F. Henry. “Social Distancing and Incarceration: Policy and Management Strategies to Reduce COVID-19 Transmission and Promote Health Equity Through Decarceration”. eng. In: *Health Education & Behavior: The Official Publication of the Society for Public Health Education* 47.4 (Aug. 2020), pp. 536–539. ISSN: 1552-6127. DOI: 10.1177/1090198120927318.
- [4] Noel A. Vest et al. “Prison population reductions and COVID-19: A latent profile analysis synthesizing recent evidence from the Texas state prison system”. en. In: *medRxiv* (Sept. 2020). Publisher: Cold Spring Harbor Laboratory Press, p. 2020.09.08.20190884. DOI: 10.1101/2020.09.08.20190884. URL: <https://www.medrxiv.org/content/10.1101/2020.09.08.20190884v1> (visited on 01/12/2021).
- [5] *uclalawcovid19behindbars/data*. original-date: 2020-06-06T00:59:09Z. Apr. 2021. URL: <https://github.com/uclalawcovid19behindbars/data> (visited on 04/27/2021).
- [6] Hao-Yuan Cheng et al. “Contact Tracing Assessment of COVID-19 Transmission Dynamics in Taiwan and Risk at Different Exposure Periods Before and After Symptom Onset”. eng. In: *JAMA internal medicine* 180.9 (Sept. 2020), pp. 1156–1163. ISSN: 2168-6114. DOI: 10.1001/jamainternmed.2020.2020.
- [7] Xi He et al. “Temporal dynamics in viral shedding and transmissibility of COVID-19”. eng. In: *Nature Medicine* 26.5 (May 2020), pp. 672–675. ISSN: 1546-170X. DOI: 10.1038/s41591-020-0869-5.
- [8] Stephen A. Lauer et al. “The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application”. In: *Annals of Internal Medicine* (Mar. 2020). ISSN: 0003-4819. DOI: 10.7326/M20-0504. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7081172/> (visited on 04/28/2021).
- [9] Andrew William Byrne et al. “Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases”. eng. In: *BMJ open* 10.8 (Aug. 2020), e039856. ISSN: 2044-6055. DOI: 10.1136/bmjopen-2020-039856.
- [10] Kieran A. Walsh et al. “The duration of infectiousness of individuals infected with SARS-CoV-2”. en. In: *Journal of Infection* 81.6 (Dec. 2020), pp. 847–856. ISSN: 0163-4453. DOI: 10.1016/j.jinf.2020.10.009. URL: <https://www.sciencedirect.com/science/article/pii/S0163445320306514> (visited on 04/29/2021).
- [11] J. O. Lloyd-Smith et al. “Superspreadering and the effect of individual variation on disease emergence”. en. In: *Nature* 438.7066 (Nov. 2005). Number: 7066 Publisher: Nature Publishing Group, pp. 355–359. ISSN: 1476-4687. DOI: 10.1038/nature04153. URL: <https://www.nature.com/articles/nature04153> (visited on 04/27/2021).

- [12] Benjamin M. Althouse et al. “Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control”. en. In: *PLOS Biology* 18.11 (Nov. 2020). Publisher: Public Library of Science, e3000897. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.3000897. URL: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000897> (visited on 04/27/2021).
- [13] Dillon C. Adam et al. “Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong”. en. In: *Nature Medicine* 26.11 (Nov. 2020). Number: 11 Publisher: Nature Publishing Group, pp. 1714–1719. ISSN: 1546-170X. DOI: 10.1038/s41591-020-1092-0. URL: <https://www.nature.com/articles/s41591-020-1092-0> (visited on 04/27/2021).
- [14] Zachary Susswein and Shweta Bansal. “Characterizing superspreading of SARS-CoV-2 : from mechanism to measurement”. eng. In: *medRxiv: The Preprint Server for Health Sciences* (Dec. 2020). DOI: 10.1101/2020.12.08.20246082.
- [15] Seth Blumberg and James O. Lloyd-Smith. “Inference of R₀ and Transmission Heterogeneity from the Size Distribution of Stuttering Chains”. en. In: *PLOS Computational Biology* 9.5 (May 2013). Publisher: Public Library of Science, e1002993. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002993. URL: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002993> (visited on 04/27/2021).

7 Supplement text: Methods

7.1 Probability of introduction

We model introduction of disease into the resident community as primarily coming from contact with infected staff. The daily rate of introducing a case, ϕ , is the product of the average number of staff-resident contacts per day and the probability that a staff-resident contact causes transmission. The average number of staff-resident contacts per day is the product of the number of susceptible individuals in the resident community, N_s , and the average number of contacts residents have with staff, N_c . Note that we define a contact to be a pairing of a resident and staff member. If there are multiple interactions of that pairing through a day, we still count it as one contact. The probability that a staff-resident contact causes transmission is the product of the prevalence of disease in the staff, P_{com} , and the probability that an infectious contact causes an infection α_{ic} . Thus,

$$\phi = N_s \cdot N_c \cdot P_{com} \cdot \alpha_{ic}. \quad (11)$$

The average number of days until an introduction occurs is the reciprocal of ϕ .

7.2 Probability of an introduction going extinct

When the number of residents with infection is a low number, the stochasticity of transmission can lead to significantly different outcomes. In some cases, introduction of disease may lead to no additional infections or a limited number of infections. In other cases, introduction of disease may lead to enough cases such that an outbreak is inevitable. We refer to the probability that an outbreak is averted as the probability of extinction, p_e .

To model the probability that introduction of disease leads to extinction, we assume that the probability distribution for the number of secondary infections caused by each new infection follows a negative binomial distribution.[11] **Lloyd-Smith Nature 2005** We parameterize the negative binomial distribution, by the reproduction number, R , and the dispersion parameter, k . The reproduction number is the average number of cases each new infection causes, but differs from R_0 in that it incorporates possible reduction of the susceptible population due to vaccination and/or depopulation. The dispersion parameter characterizes the degree of transmission heterogeneity. Low values of k are seen for disease in which superspreading occurs, meaning that a relatively few number of cases causes a large proportion of the cases.

With these assumptions, there is an analytic relationship between R , k , the number of disease introductions occurring concurrently, ζ , and the probability, $q_s(N_i)$, of having an outbreak size of s . [15] **Ref Blumberg PLoS CB** The

probability, $r_{i,j}$, that i infections causes j infections is,

$$r_j(i) = \frac{\Gamma(j+k \cdot i)}{\Gamma(j+1) \cdot \Gamma(k \cdot i)} \cdot \left(\frac{k}{R+k} \right)^{k \cdot i} \cdot \left(\frac{R}{R+k} \right)^j, \quad (12)$$

where Γ is the Gamma function. Then,

$$q_s(\zeta) = \frac{\zeta}{s} \cdot r_{s-\zeta}(s). \quad (13)$$

The use of a stationary negative binomial offspring distribution assumes that depletion of susceptibles is not a significant factor in transmission dynamics. This is a reasonable assumption if the number of infections that follow one or more introductions is small. However, once there are a significant number of infections, C_i , following one or more introductions, a different transmission model is needed that accounts for depletion of susceptibles. We classify those instances in which there are more than C_{th} cases as an outbreak. All clusters of infection arising from an introduction that have less than C_{th} cases are considered extinction events. Thus the probability of a disease introduction leading to extinction is,

$$p_e = \sum_{s=\zeta}^{C_{th}} q_s(\zeta). \quad (14)$$

7.3 Size and temporal dynamics of an outbreak

Once an outbreak is established, we use a deterministic compartmental model to describe the transmission dynamics. In this model, the rate that the susceptible residents become infected is proportional to the product of the number of susceptible and infectious residents. Once infected, a resident will be classified as exposed, but not infectious. The rate that exposed residents transition into being infectious is proportional number of residents who are in the exposed state. Infectious residents are removed from the population at a constant rate. Removal may occur due to recovery, transport for higher level of care or death. Once residents are removed, they are considered non-infectious and are no longer susceptible. During the course of an outbreak, it is assumed that no new susceptible residents are introduced into the population. The equations describing the model are,

$$\frac{dS}{dt} = -\frac{S}{N} \cdot \frac{R \cdot I}{T_I} \quad (15)$$

$$\frac{dE}{dt} = \frac{S}{N} \cdot \frac{R \cdot I}{T_I} - \frac{E}{T_E} \quad (16)$$

$$\frac{dI}{dt} = \frac{E}{T_E} - \frac{I}{T_I} \quad (17)$$

$$\frac{dX}{dt} = \frac{I}{T_I} \quad (18)$$

where S , E , I , and X represent the number of susceptible, exposed, infectious, and removed residents. The average time for being in the exposed and infectious states is represented by T_E and T_I respectively.

To determine the total number of residents infected, I_{tot} , the proportion of residents who become infected, P_T , the maximum number of residents infected at once, M_T , and the time of peak incidence, T_M , we simulate the preceding model numerically. Then,

$$I_{tot} = X(t \rightarrow \infty) \quad (19)$$

$$P_T = \frac{I_{tot}}{N} \quad (20)$$

$$M_T = \max_{t \in (-\infty)} I(t) \quad (21)$$

$$T_M = I(t = M_T). \quad (22)$$

All calculations and simulations are conducted in R.

8 Supplemental Figures

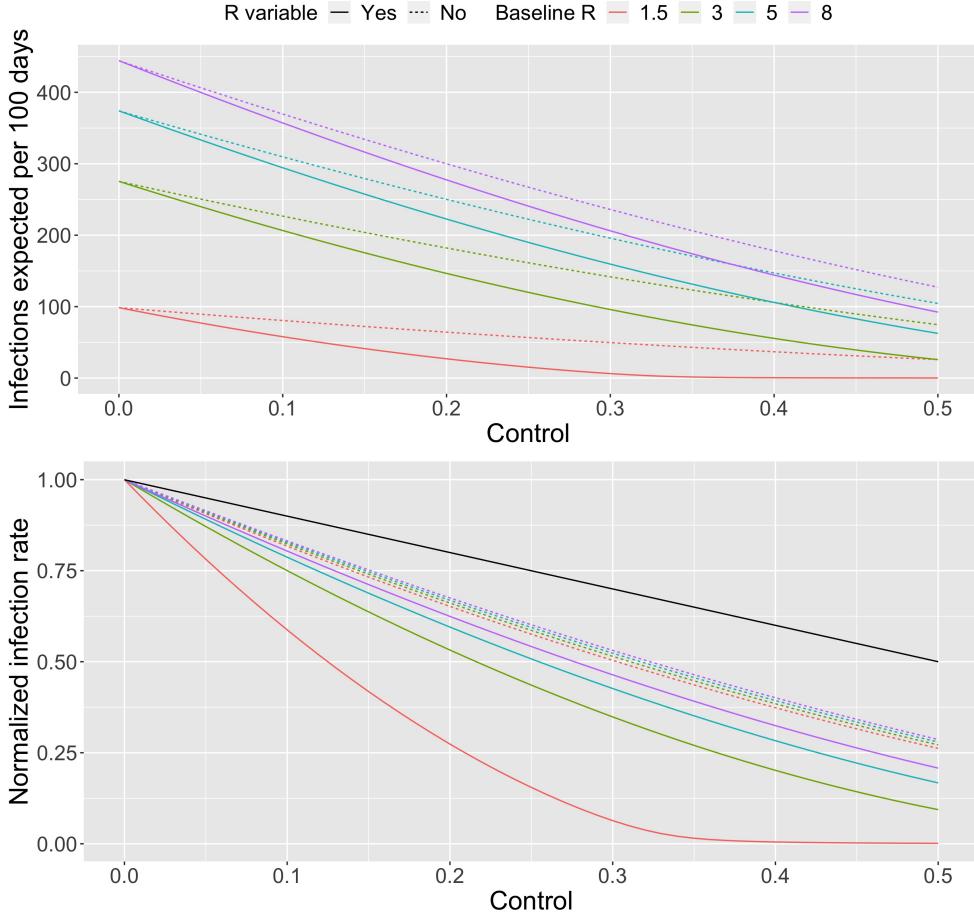


Figure 3: Impact of preemptively decreasing the size of the susceptible population. (Top) The expected number of infections due to outbreaks occurring within 100 days as a function of the level of control. The control value shown on the x-axis is the proportion of the number of susceptible individuals that are removed from the resident population via vaccination and/or depopulation. Colors correspond to different values of the baseline reproduction number, R_0 , for when no control is implemented. Specific values of the baseline R_0 are specified by the legend. Dashed lines show the result assuming that the control intervention does not change the reproduction number (e.g. $\Theta = 0$ corresponding to a vaccine that immunizes against disease, but not asymptomatic infection). Solid lines show results assuming that the reproduction number changes in proportion to the level of control (e.g. $\Theta = 1$ corresponding to a vaccine that immunizes against disease and asymptomatic infection). (Bottom) Analogous to top panel, except that the rate of infection has been normalized to a rate of one when the level of control is zero. This highlights the relative impact of decreasing the number of susceptible individuals. For both panels superspreading is allowed, which is modeled by a dispersion value of 0.2. We set $N_C = 10$ and $\alpha_{ic} = 0.01$, meaning that residents are assumed to have 10 contacts with staff a day and that contact with an infected staff has a one percent chance of transmitting infection. The other parameters are the same as the defaults for Figures 6 – 9.

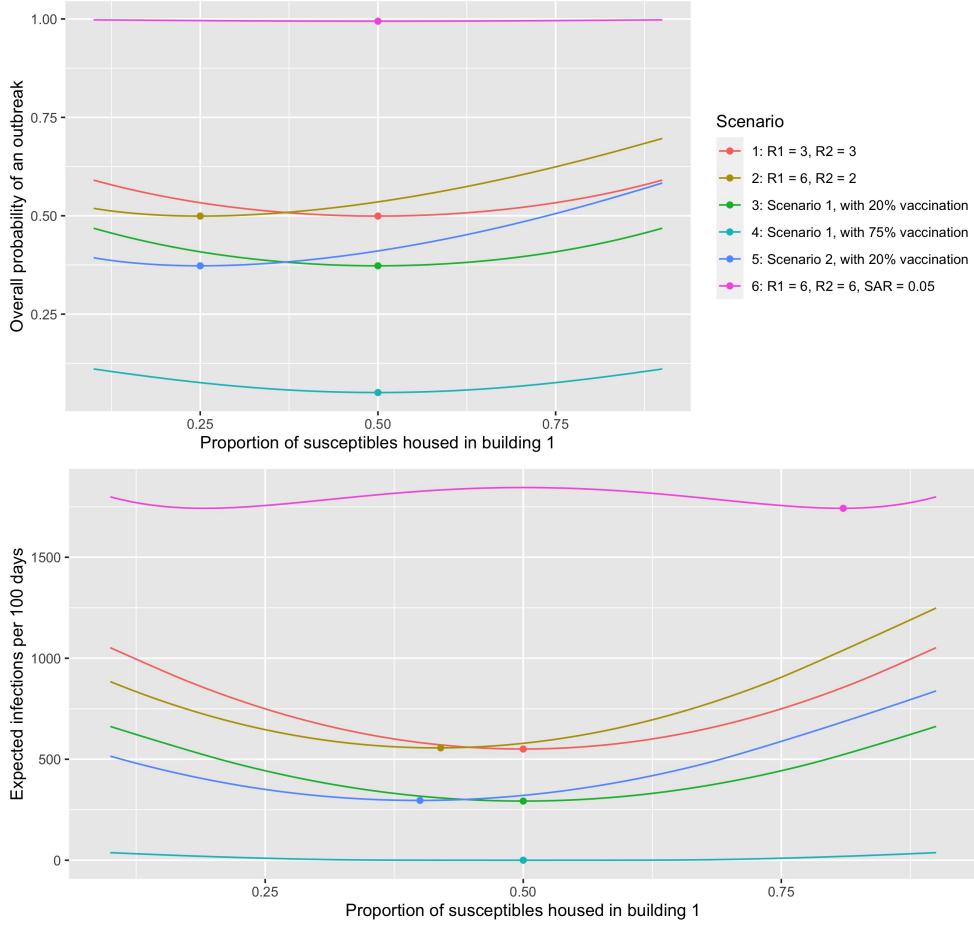


Figure 4: Optimal distribution of residents. (Top) Probability that an outbreak occurs in at least one of two buildings within 100 days as a proportion of the total population that is housed in Building 1. The total number of residents is assumed to be 2,000 so that a proportion of residents housed in building 1 of 0.25 corresponds to 500 residents in building 1 and 1,500 in building two. The colored lines represent six different scenarios as indicated by the legend. The reproduction values in the legend correspond to the reproduction number if residents are distributed equally in both buildings. In scenario 1 the transmission in the two buildings are identical. In scenario 2, building 1 has higher transmission than building 2. Scenarios 3, 4 and 5, are identical to scenarios 1 and 2 except that a fraction of the residents in each building are vaccinated. In scenario 6, which is designed to model a very risky housing environment, the probability that an infected staff transmits infection to a resident is increased five-fold compared to the other scenarios. (Bottom) The expected number of infections per 100 days for both buildings combined is shown as a function of the proportion of residents in building 1. The bottom panel is otherwise analogous to the top panel. Except as noted above, the parameter values are identical to figure 3. Will align the two plots vertically so points along the x axis can be compared

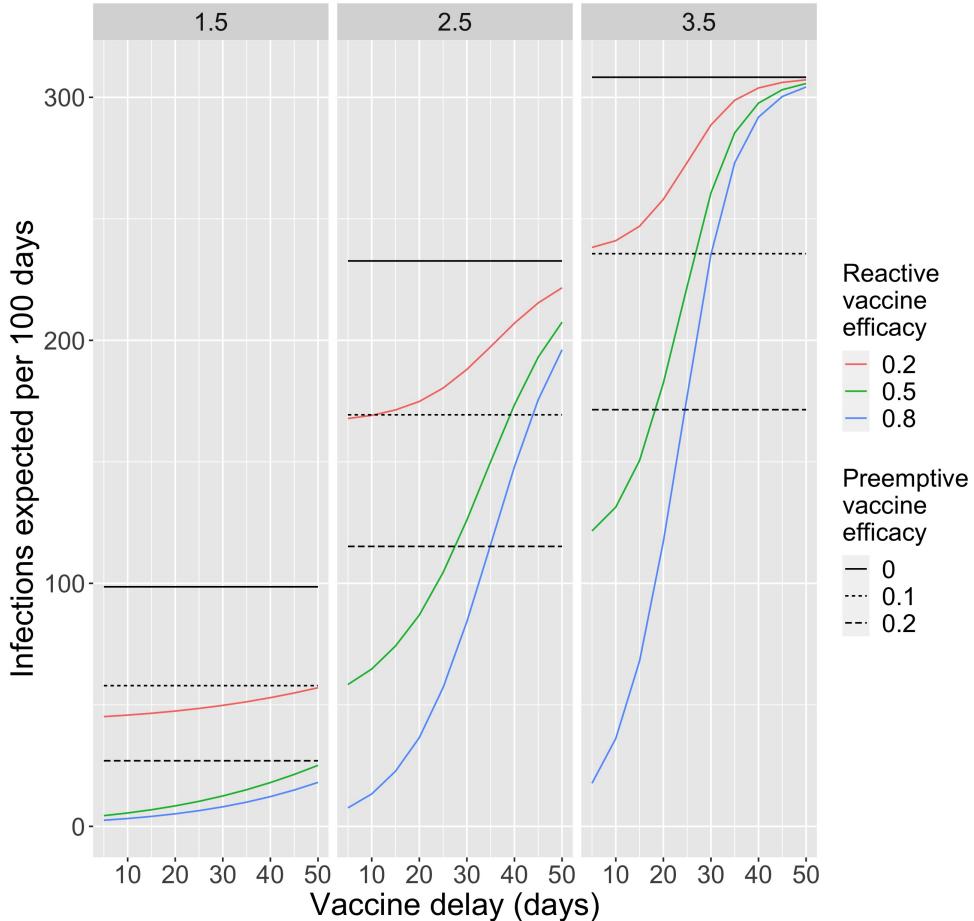


Figure 5: Preemptive vs reactive vaccination The number of infections per 100 days is plotted for various vaccine scenarios. Each vertical panel corresponds to a different value of R_0 as depicted on the top bar. Horizontal black bars correspond to no vaccination being administered. Horizontal dashed bars correspond to preemptive vaccination with vaccine efficacy indicated in the lower legend. Colored curves indicate the expected number of infections when vaccine is given reactively once ten cases in a building are identified. We model the vaccine as being effective at reducing transmission after a predetermined vaccine delay. The delay is depicted on the x-axis and a vaccine efficacy as depicted by the upper legend. Here, the vaccine efficacy is used to quantify the combined impact of vaccine coverage and the ability of the vaccine to prevent transmission (E.g. if two third of residents accept the vaccine and the vaccine decreases transmission by 75% then the vaccine efficacy is 50%). Parameter values are otherwise identical to those in Figure 3.

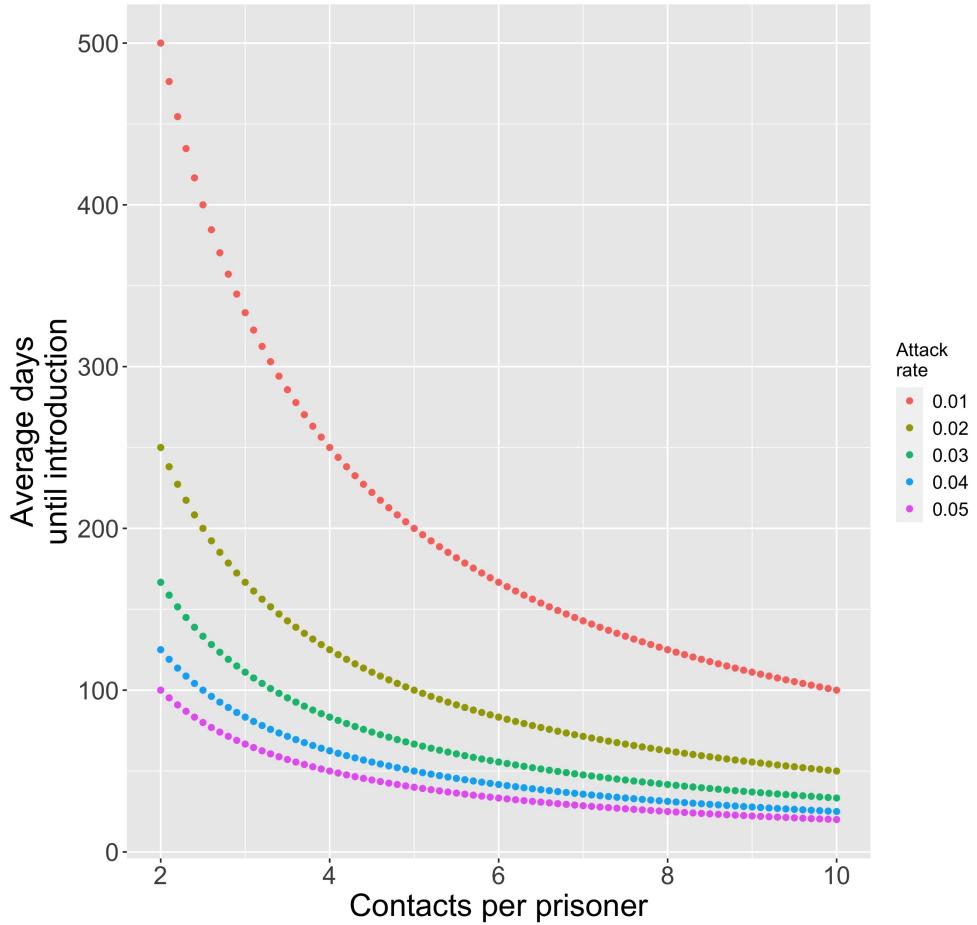


Figure 6: Frequency of an introduction. Average number of days until a COVID case is introduced into the resident community, ϕ , as a function of the average number of contacts a resident has with resident staff, N_c . Colors correspond to different values for α_{ic} , the probability that a resident's contact with an infected staff causes an infection. The prevalence of disease in the community, P_{com} , is assumed to be 0.01%. The number of susceptible individuals in the congregate community, N_s , is assumed to be 1,000.

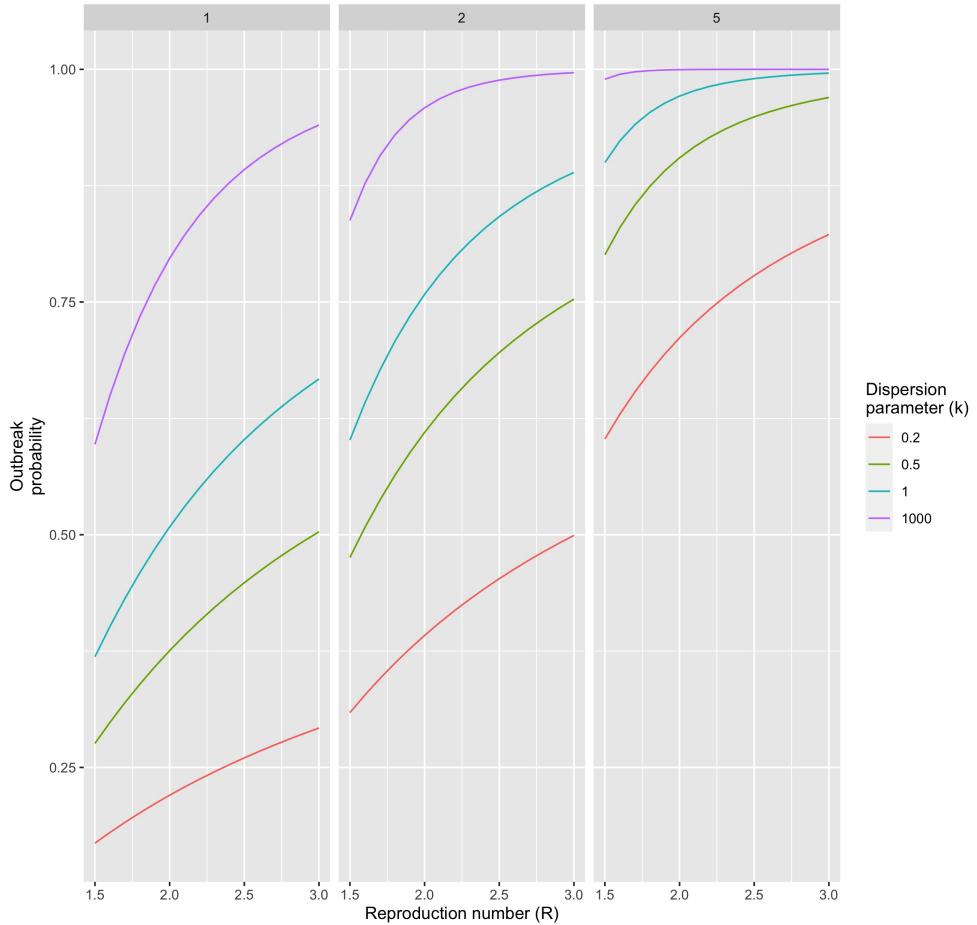


Figure 7: Probability of an outbreak. The probability of an outbreak occurring as a function of the reproduction number, which is defined the average number of transmission events each new cases causes. The panels correspond to different numbers of simultaneous disease introductions as indicated in the top of the panel. The different colors correspond to different values of the dispersion parameter. Homogeneous transmission corresponds to $k = \infty$, and superspread ing is more prevalent as k decreases. Plots are based on $C_{th} = 10$, meaning that an outbreak is defined to occur when an introduction leads to at least ten cases.

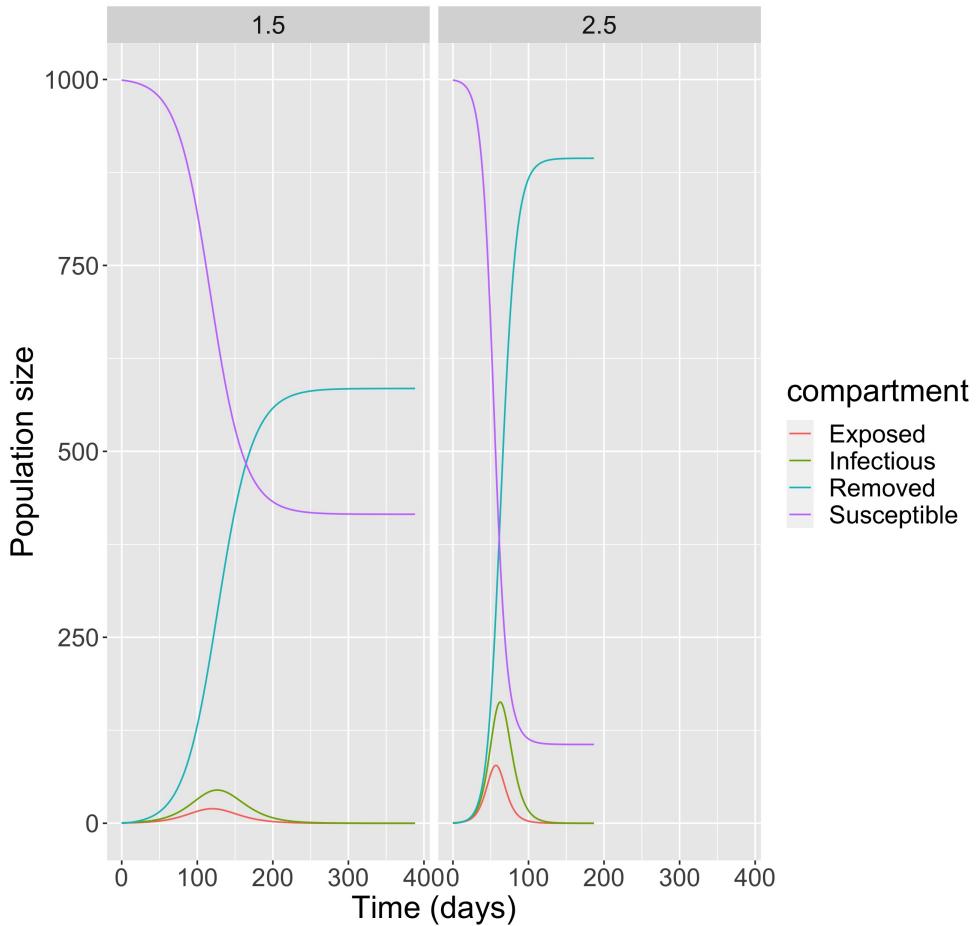


Figure 8: **Transmission dynamics.** The number of susceptible, exposed, infectious and removed residents are shown as a function of time for an outbreak in a congregate facility. The removed category includes those who have recovered from illness, those who are sick but quarantined and those who have died. The two panels represent two different values of R . A total population of 1,000 is assumed. The average duration of each covid case being in the incubation and infectious periods is 3 and 7 days respectively. The depicted outbreaks start with one exposed individual at time 0. The time step used for running the transmission dynamics model is 0.2 days, and the model is run until a negligible number of infectious individuals remain.

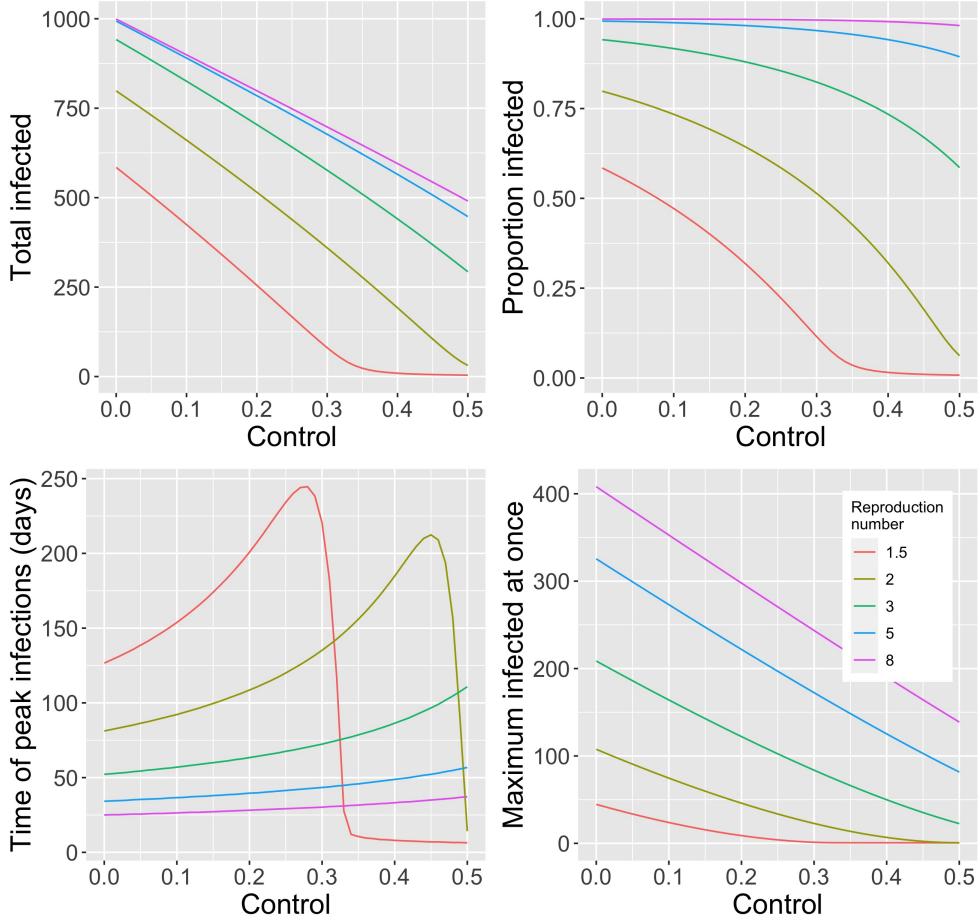


Figure 9: **Outbreak size and timing.** Each panel shows a feature of outbreak dynamics as a function of the control factor, which is the proportion of the population that is no longer susceptible due to depopulation or vaccination. The different colored lines correspond to different values of R_0 when there is no decrease in the size of the susceptible population. The panels show the impact on the total number of residents who are infected (upper left), the proportion of susceptible residents who become infected (upper right), the timing of the peak in the number of infectious individuals (lower left), and the maximum number who are infectious at once (lower right). In all panels it is assumed that the effective reproduction number scales linearly with the control factor. Parameter values are the same as for figure . For lower values of baseline R_0 the timing of the peak in the number of infectious individuals abruptly decreases when the control factor is sufficiently high such that the effective reproduction number is less than one and no significant peak in infectious cases occurs (e.g. when control is about 30% for as baseline $R_0 = 1.5$ as seen in the lower right panel.)

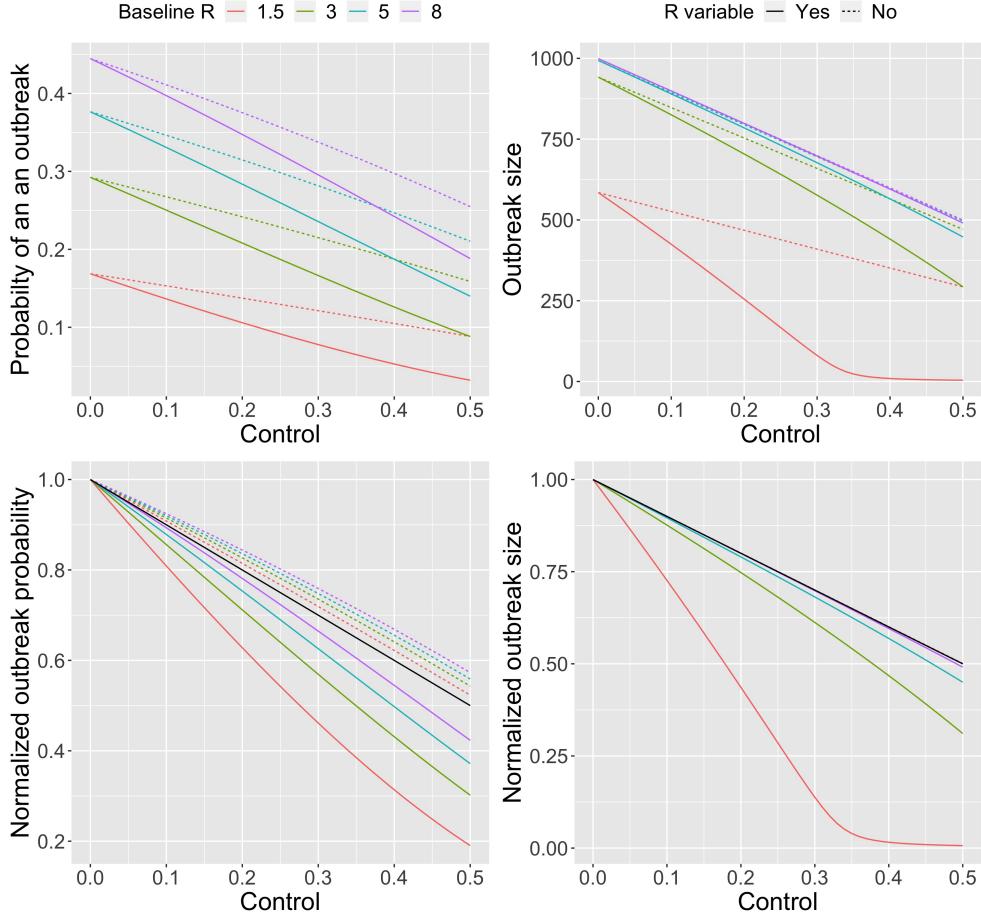


Figure 10: Probability and size of an outbreak Probability of an outbreak occurring in 100 days (left panels) and size (right panels) of outbreaks as a function of control, defined as the proportion of the susceptible population reduced by vaccination and/or depopulation. As parameterized by the legend, colors correspond to different values of the baseline reproduction number that represents transmission when there is no control. Top panels show the absolute values, while the bottom panels show normalized values that are always one when there is no control. Parameter values are as specified in Figure 3. As a technical note, there is a less than linear impact on the probability of an outbreak when control does not affect the reproduction number (dashed lines in lower left panel), because for our chosen parameter values the probability of an outbreak begins to saturate when the level of control is very small.

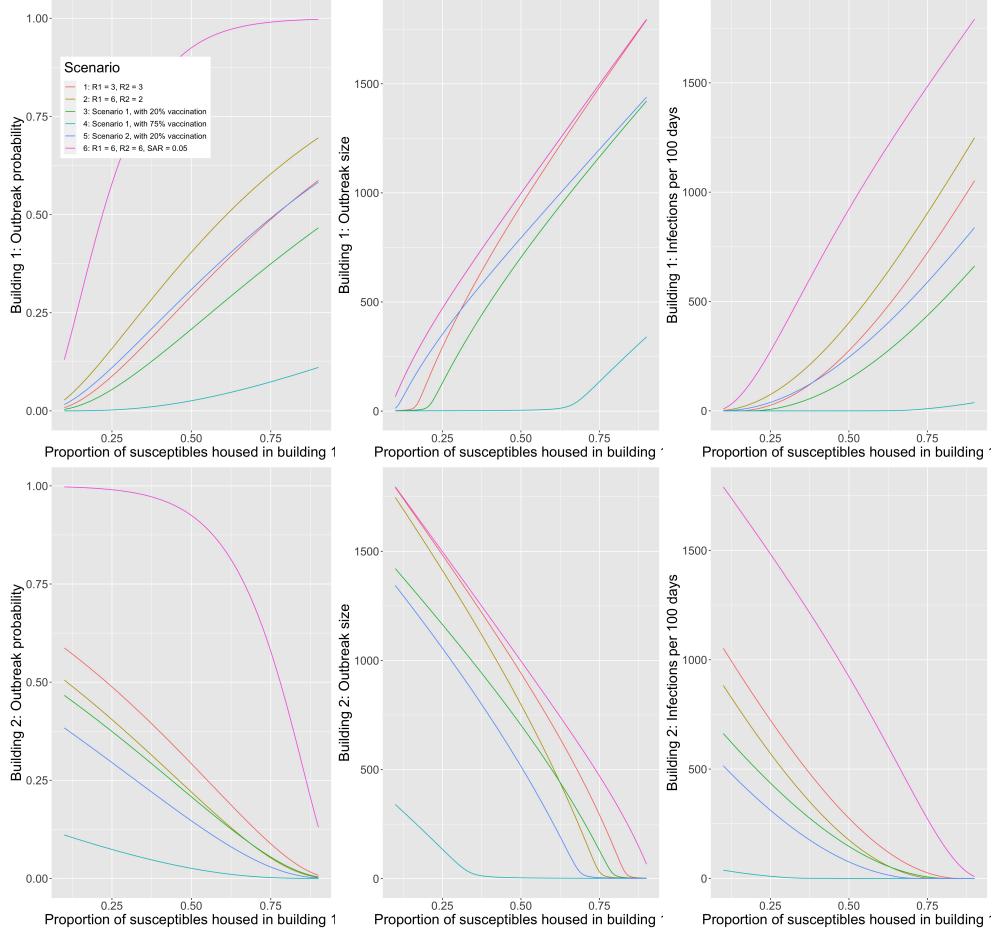


Figure 11: Outbreak dynamics depend on how residents are distributed between two buildings. The probability of an outbreak occurring in 100 days (left panels), expected outbreak size (middle panels) and the overall expected number of cases in 100 days (right panels) are shown as a function of the proportion of residents housed in the first of two buildings. Top panels corresponds to building 1 and the bottom panels correspond to building 2. Parameter values and scenarios are the same as for Figure 11. *Need to fix x-axis label*