UC Berkeley

UC Berkeley Previously Published Works

Title

Evaluating strategies for spatial allocation of vaccines based on risk and centrality.

Permalink

https://escholarship.org/uc/item/7vg3n57q

Journal

Journal of the Royal Society Interface, 19(187)

Authors

Singer, Benjamin

Thompson, Robin

Bonsall, Michael

Publication Date

2022-02-01

DOI

10.1098/rsif.2021.0709

Copyright Information

Peer reviewed

THE ROYAL SOCIETY
PUBLISHING

Evaluating strategies for spatial allocation of vaccines based on risk and centrality

Benjamin J. Singer[1], Robin N. Thompson[2,3] and Michael B. Bonsall[1]

[1]Department of Zoology, University of Oxford, Oxford, UK
[2]Mathematics Institute, and [3]Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Coventry, UK

BJS, 0000-0003-3062-2620; RNT, 0000-0001-8545-5212; MBB, 0000-0003-0250-0423

When vaccinating a large population in response to an invading pathogen, it is often necessary to prioritize some individuals to be vaccinated first. One way to do this is to choose individuals to vaccinate based on their location. Methods for this prioritization include strategies that target those regions most at risk of importing the pathogen, and strategies that target regions with high centrality on the travel network. We use a simple infectious disease epidemic model to compare a risk-targeting strategy to two different centrality-targeting strategies based on betweenness centrality and random walk percolation centrality, respectively. We find that the relative effectiveness of these strategies in reducing the total number of infections varies with the basic reproduction number of the pathogen, travel rates, structure of the travel network and vaccine availability. We conclude that when a pathogen has high spreading capacity, or when vaccine availability is limited, centrality-targeting strategies should be considered as an alternative to the more commonly used risk-targeting strategies.

1. Introduction

Administrators of vaccination programmes often find that there are fewer doses of vaccines available (or capacity to administer these doses) than there are susceptible individuals willing to receive them. In this situation, some form of prioritization is necessary, in which certain individuals are chosen from the susceptible population to be vaccinated, or to receive vaccines earlier than others. Often, vaccines are offered to individuals in certain demographics who are particularly at risk of catching, spreading or experiencing severe outcomes of a disease. This kind of demographic targeting is used by the United Kingdom's National Health Service to distribute influenza vaccines—the NHS offers free vaccines to individuals over 65, or with conditions such as asthma and diabetes, who are at risk of severe influenza [1]. Targeting demographics at high risk of severe disease has also been a key strategy in COVID-19 vaccination programmes, with elderly people often being prioritized [2–4].

An alternative way of selecting which individuals receive vaccinations is by virtue of their location in geographical space. Spatial position can be used as an alternative to demographic membership as a predictive variable for risk of infection [5], but often spatial strategies are used when cases of a disease occur only within a certain region, and public health decision-makers seek to contain the pathogen locally and prevent it from spreading to other regions [6–8]. Spatial control can be particularly important when responding to pandemic risks at the point when containment is still possible [9]. While we focus on vaccination to provide a concrete example of a spatial control tool, the structure of our model is general and can represent any prophylactic control strategy that prevents infections (e.g. administering infection-preventing drugs, or culling of animals and inducing resistance in plants in the case of agricultural diseases).

Within spatial strategies, there are many ways in which certain regions can be prioritized for vaccine distribution. Among alternatives, priority could be given

to regions close to an existing epidemic, with high population density, or with many travel connections. In practice, many modelling studies that consider spatial vaccine allocation use a risk-targeting approach, in which those regions predicted to be most at risk of importation of a pathogen are prioritized for distribution of the vaccine [5,8,10]. There is, however, no proof that this leads to the maximum reduction in infections, and authors often mention possible further work examining 'downstream effects' [10] or 'spreading capacity' [11] of certain regions, or the 'impact of vaccination on … spatial expansion' [5].

Here we define a vaccine strategy as a method for ordering all individuals susceptible to an infection from highest priority to be vaccinated, to lowest priority. We define vaccine allocation as a division of those individuals into two sets—those in one set receive a dose of vaccine, and those in the other do not. We define risk-targeting as any vaccine strategy in which a region is prioritized for receipt of vaccine doses based solely on the risk that a large outbreak of the relevant pathogen will occur in that region. The method for determining this risk can range from simple heuristics to sophisticated epidemiological forecasts [8]. One particularly widely discussed form of risk-targeting is spatial ring vaccination (as opposed to ring vaccination of contacts), in which individuals are prioritized for vaccination based on their proximity to an ongoing outbreak. This kind of strategy has been discussed in settings from foot-and-mouth disease in the United Kingdom to Ebola in the Democratic Republic of the Congo [6,7].

An alternative to risk-targeting in strategies for vaccine distribution is centrality-targeting [12–14]. Centrality is a term for any measure of the importance of a node in a network. One possible use of some centrality measures is as proxy for the spreading capacity of a given region. We discuss specific measures of centrality, and their epidemiological relevance, in more detail below. In centrality-targeting strategies, vaccines are preferentially allocated to regions with high centrality in a travel network. Thus, they may act to impede the circulation of a pathogen between regions, rather than directly protecting those regions most likely to experience outbreaks themselves.

There is little existing work directly comparing risk-targeting of vaccines to centrality-targeting. A study by Piraveenan *et al.* [12] compares targeting of prophylactic control in a disease transmission network according to three different measures: hop distance (a simple kind of risk-targeting), percolation centrality and betweenness centrality [12]. Hop distance is the length of the shortest path through the network from a susceptible node to the closest infected node, and can be treated as a proxy for risk. Betweenness centrality is a very commonly used centrality measure based on a node's connectivity. Percolation centrality is a measure of node centrality introduced in the above study, which takes into account a node's connectivity as well as the current state of infection in the network. Piraveenan *et al.* find that, in a scale-free and a random network, which of the three strategies best reduces infections depends on parameters reflecting how widespread the pathogen is, and how many nodes can be immunized. In general, when the pathogen is scarce in the network, or when many nodes can be immunized, risk-targeting (that is, priority according to hop distance), is most effective at reducing infection. In the opposite scenario, when the pathogen is widespread or when few nodes can be immunized, betweenness-targeting is most effective. In the intermediate regime, targeting according to percolation centrality led to the largest expected reduction in infections.

Generalizing from this result, we conjecture that risk-targeting may be less effective than centrality-targeting when resources available for control of a pathogen are very limited (e.g. few doses of a vaccine), or when the pathogen concerned has a high spreading potential (either by having a high R_0 or by travelling widely across space). We argue that Piraveenan *et al.*'s findings are intuitive—for instance, when many resources are available for control, a pathogen can be fully contained by ring vaccination (a form of risk-targeting), while the only chance to contain a pathogen when few resources are available is by targeting travel hubs (a form of centrality targeting). This intuitive understanding of the scenario leads to the following generalized hypothesis:

— when resources for control are widely available, and the potential for widespread pathogen transmission is limited, targeting regions at high risk is more effective than targeting regions with high centrality.
— when resources are limited and the pathogen has the potential to spread widely, targeting regions with high centrality is more effective than targeting regions at high risk.

To investigate this hypothesis, we examine whether the risk-targeting strategy is the most effective strategy to reduce the overall number of infections in a range of different epidemiological networks. We compare this strategy to two variations on a centrality-targeting strategy that prioritizes regions for vaccination according to their centrality on the human travel network. The variations come about due to alternative centrality measures.

2. Methods

2.1. Disease model

In this study, we make use of a simplified stochastic SIR metapopulation model similar to others used in spatial analyses of infectious disease epidemics and their control [5]. Each node i on the travel network is associated with a region, which has its own population size N_i and within-region basic reproduction number $R_{0,i}$. We can represent the travel network with a matrix λ, in which the element λ_{ij} is the expected number of individuals travelling from region i to region j per unit time. Like all rates referred to in this section, this is the rate of a stochastic process, where the actual number of individuals travelling over a time period Δt is given by a Poisson distribution with rate parameter $\lambda_{ij}\Delta t$.

With this information, we can calculate the expected number of infected individuals travelling from a region i that experiences an epidemic, to any other region connected to it on the travel network. First we calculate the total number of infected individuals over the course of the epidemic in region i, $R_i(\infty)$, using the widely used SIR final size equation [15–19]:

$$R_i(\infty) = N_i - N_i \exp\left(\frac{-R_{0,i}R_i(\infty)}{N_i}\right). \tag{2.1}$$

Each infected individual can either recover while still in region i, or travel to a neighbouring region before recovering. The probability P_{ij} that an individual travels to a particular neighbouring region j while infected is therefore

$$P_{ij} = \frac{\lambda_{ij}}{\sum_k \lambda_{ik} + \mu}, \tag{2.2}$$

where μ is the recovery rate, and $\sum_k \lambda_{ik}$ is the total rate of travel out of region i. This means that the expected total number I_{ij} of infected

individuals who arrive in region j from region i over the course of the epidemic in i is

$$I_{ij} = \frac{\lambda_{ij} R_i(\infty)}{\sum_k \lambda_{ik} + \mu}. \tag{2.3}$$

When $\mu \gg \sum_k \lambda_{ik}$, meaning that individuals are much more likely to recover than travel while ill, equation (2.3) can be approximated as:

$$I_{ij} \approx \frac{\lambda_{ij}}{\mu} R_i(\infty). \tag{2.4}$$

This approximation allows us to treat disease dynamics separately in each region, with the course of a local epidemic progressing independently of its neighbours once it has been seeded. We also make use of this approximation to assume that the population size N_i of each region does not change over the course of the outbreak.

To avoid highly intensive computations, let us consider a simplified disease model with a square epidemic curve—meaning that, in our model, there are a constant number of infected individuals during an epidemic and no infected individuals otherwise. We assume that the duration of such an epidemic is exponentially distributed, but that the number of infected individuals is deterministic and independent of the epidemic duration. This assumption of the exponential distribution of epidemic duration is the greatest departure in the behaviour of our simplified model from that of the full stochastic SIR model. We discuss the implications of this further in the electronic supplementary material. Since the size of the epidemic does not vary from its start to its end, this means that we can shift from treating the states of individuals, to treating the states of whole regions. We can treat each region as belonging to a state S, meaning that no epidemic has occurred; I, meaning that an epidemic is ongoing; or R, meaning that an epidemic has occurred and the region is protected from further outbreaks due to herd immunity.

The expected total number of infected individuals who arrive in region j from a region i, which experiences a square epidemic curve with height A and mean duration D, is given by $\lambda_{ij} AD$. This quantity determines the probability for the pathogen to spread between regions, so we seek to match its value in our simplified model to the corresponding value in equation (2.4). To achieve this, we choose $A = R_i(\infty)/\mu$ and $D = 1$. This choice ensures that one unit of time is equal to the expected duration of a single epidemic, thus effectively rescaling time to match the simplified pathogen dynamics.

In a stochastic SIR model, the probability for a single infected individual to cause an epidemic when introduced into a large, entirely susceptible population is $1 - 1/R_0$, a result cited widely in the epidemiology literature [16,20–24]. So if the rate of arrival of infected individuals in region j from region i is $\lambda_{ij} R_i(\infty)/\mu$, and each arrival has a probability of causing an epidemic given by $1 - 1/R_0$, then the rate at which the epidemic in region i seeds an epidemic in region j is

$$\left(1 - \frac{1}{R_0}\right)\lambda_{ij} R_i(\infty)/\mu. \tag{2.5}$$

This completes the structure of our model: regions are assigned states S, I and R; regions in state I export the epidemic stochastically at the above rate to neighbouring regions in state S; and regions in state I recover stochastically into state R at rate 1, after which they cannot be reinfected. We assume that the time between events is exponentially distributed, and use Gillespie's direct stochastic simulation method to generate our results [25]. To characterize the behaviour of an outbreak, we run many simulations using this model and take the mean of the total cumulative number of infections at the end of each simulation.

2.2. Centrality measures

The alternatives to risk-targeting that we investigate are centrality-targeting strategies. Centrality is a measure of the importance of a node in a network. Targeting nodes for intervention based on their centrality may help prevent wide circulation of a pathogen in a transmission network. There are several different ways of measuring centrality, including degree centrality, eigenvector centrality, closeness centrality and many others. One particularly widespread measure of centrality is betweenness centrality.

The betweenness centrality of a node v is the proportion of shortest paths from a source node s to a target node r that pass through v, averaged over all sources and targets s and r [26,27]. In a weighted travel network, the contribution to the length of a path of one of its constituent edges is equal to the reciprocal of the weight of that edge, since higher weights represent stronger connections. The betweenness b_v of node v in a network G of size n is expressed as

$$b_v = \frac{1}{(n-1)(n-2)} \sum_{s \neq v, r \neq v} \frac{g_v^{(s,r)}}{g^{(s,r)}}, \tag{2.6}$$

where $g^{(s,r)}$ is the number of geodesics (shortest paths) from node s to node r, and $g_v^{(s,r)}$ is the number of such paths that pass through node v. This definition ensures that the betweenness of a given node is always between 0 and 1.

Betweenness centrality is discussed widely in studies of spreading processes on networks [28–31]. The measure's definition in terms of paths through a network makes it more appropriate for use in the context of spreading processes than, say, degree centrality [32]. For this reason, we include it in our analysis (see the electronic supplementary material for a comparison of betweenness centrality to degree centrality). However, there is good reason to believe that betweenness centrality is not particularly well suited to finding important nodes in a disease transmission network. First, it only counts geodesic paths—but infections transmitting between regions cannot be assumed to follow the shortest path from their source to their destination. Second, when responding to an invading pathogen, we often know the possible sources of an infection and are seeking to disrupt transmission from these sources—whereas the betweenness centrality calculation averages across all possible sources. We can resolve the first issue with random walk betweenness centrality, and the second issue with percolation centrality—combining the two results in our random walk percolation centrality (RWPC) measure.

Random walk betweenness centrality is defined similarly to betweenness centrality, but expands the paths counted, from only shortest paths from node s to node r that pass through node v, to all random walks that pass through v from s to r [33,34]. Since a random walk can have an arbitrarily high number of steps, we allow steps passing along an edge in opposite directions to cancel out, so that we ignore random walks that go back and forth through a given node many times without progressing to another part of the network. We can then define the random walk betweenness b_v^{rw} of a node v in terms of the net number of times a random walk from s to r is expected to pass through v, $I_v^{(s,r)}$.

$$b_v^{rw} = \frac{1}{(n-1)(n-2)} \sum_{s \neq v, r \neq v} I_v^{(s,r)}, \tag{2.7}$$

The method for calculating $I_v^{(s,r)}$ is given in Newman [33].

Percolation centrality introduces a weighting to each element of the sums in equations (2.6) and (2.7) depending on the infection state of the source node s [12]. The percolation centrality $p_v(t)$ of node v at time t is

$$p_v(t) = \frac{1}{(n-2)} \sum_{s \neq v, r \neq v} \frac{g_v^{(s,r)}}{g^{(s,r)}} \frac{x_s(t)}{[\sum_i x_i(t)] - x_v(t)}, \tag{2.8}$$

where $x_i(t)$ is a scalar quantity representing the infection state of node i at time t. In our analysis, we assume that a given region is either infected (state I) or uninfected (state S or R), restricting the
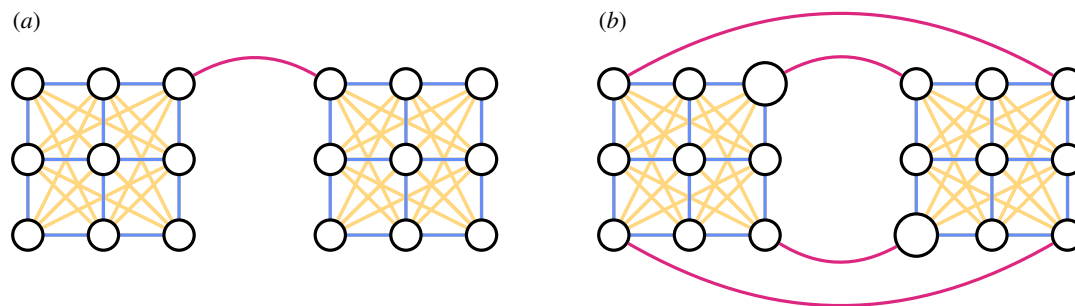
Figure 1. The 'joined grids' arrangements of regions (circles), in which the movement of the population happens between adjacent regions in each grid (blue), on longer-range travel within a grid (yellow), or along inter-grid travel routes (red). The area of the circles scales with the population of the corresponding region. (*a*) The 'one-to-one' joined grids network. (*b*) The 'four corners' variation.

values of $x_i(t)$ to 0 if region i is in state S or state R, and 1 if region i is in state I. In this binary case, equation (2.8) can be interpreted as equivalent to equation (2.6), with the source node s restricted to the class of infected nodes, and the quantity normalized appropriately. This successfully captures the importance of paths starting from an infected node in a disease transmission network.

Combining the modifications to betweenness centrality introduced in equations (2.7) and 2.8 gives us RWPC. The RWPC $p_v^{rw}(t)$ of a node v at time t is

$$p_v^{rw}(t) = \frac{1}{(n-2)} \sum_{s \neq v, r \neq v} I_v^{(s,r)} \frac{x_s(t)}{[\sum_i x_i(t)] - x_v(t)}. \quad (2.9)$$

Note that this equation is to equation (2.7) as equation (2.8) is to equation (2.6).

To apply this measure to human mobility networks, we need to capture the fact that the edges of these networks are weighted by the amount of travel that occurs along them, and can be directed if travel is not symmetric between regions. In the electronic supplementary material, we present the novel mathematics of RWPC in a weighted directed network. This results in the measure of RWPC used in this study.

2.3. Network specifications

We tested vaccination targeting strategies for simulated epidemics on metapopulations with a variety of transport networks. We chose four network structures, two artificial and two empirical. Our artificial networks are based on a 'joined grids' structure, in which the regions are split into two identical subsets, each of which is arranged in a grid with unit distance between orthogonally adjacent regions (figure 1). Travel occurs at the highest rate between regions (represented by nodes) which are adjacent on the grid, and at a rate λ_{ij} determined by distance between more distant regions. In particular,

$$\lambda_{ij} = \frac{\lambda}{(r_{ij})^a}, \quad (2.10)$$

where λ is the travel rate between adjacent regions, $r_{ij} \geq 1$ is the distance between region i and region j, and a is a parameter determining the amount of long-distance travel—the higher the value of a, the less long-distance travel. This is a kind of gravity model, which is one of the standard models of human movement [35–38]. In the basic version of this network, illustrated in figure 1*a*, there is a single travel route between the two grids, which runs from a region at the corner of one grid to a region at the corner of the another. We set the travel rate along this route at $\lambda_{\text{join}} = \lambda$ by default. With this model, we can represent outbreaks occurring at three different scales—within regions (represented by circles in figure 1), within grids (each grid is made up of nine regions) and across the whole network (the whole network is made up of 18 regions on two grids). We assume that each region has a fixed population of 200 000.

To explore further the role of population size and alternate routes of transmission in an artificial network, we created a variation on the basic joined grids arrangement, which we call the 'four corners' network. In this network, the travel within each grid is given by equation (2.10), but there are four routes between the two networks, joining each corner of each grid to a corner of the other grid, as illustrated in figure 1*b*. The inter-grid routes each have travel rate $\lambda_{\text{join}} = \lambda/4$. Each region has a fixed population of 200 000, except for one region on the corner on each grid, which has a population of 400 000, to represent high-population travel hubs. These two regions are not directly connected to each other. To account for the greater travel to more populous regions predicted by the gravity model of human movement, we now write

$$\lambda_{ij} = \frac{\lambda N_i}{(r_{ij})^a N}, \quad (2.11)$$

for the within-grid travel rates, where $N = 200\,000$.

To extend our theoretical results to real-world networks, we make use of two freely available data sets concerning human movement patterns. The first is a subset of the US domestic air traffic network, sourced from the Bureau of Transport Statistics [39]. We used as nodes in our network only the 20 airports with the highest total sum of passengers arriving and passengers departing, and defined the weight of a directed edge between two of these nodes i and j in proportion to the total number of passengers travelling from airport i to airport j in 2019. To assign a population size to each node, we used the population size of the urban area in which the airport is situated [40]. There are two airports in the data set situated in Chicago. In this case, we merge the two airports, assigning the population of the Chicago urban area to the resultant node and including all passengers travelling to and from either Chicago airport when calculating the weights of edges involving the resultant node.

The edge weight λ_{ij} (i.e. the travel rate) from node i to node j in this network is therefore given by

$$\lambda_{ij} = s \frac{t_{ij}}{N_i}, \quad (2.12)$$

where t_{ij} is the total number of passengers travelling from the airports(s) represented by node i to the airport(s) represented by node j in 2019, $N_i > 0$ is the population size of node i, and s is an overall scaling factor that we allow to vary.

The second empirical network we use is sourced from the 2011 UK census [41]. We use the census to estimate the amount of travel between all cities and towns with populations over 100 000 in an area of Northwest England including Greater Manchester and West and South Yorkshire (figure 2). The census includes data for the location of the residence and place of work for all respondents, which we use to estimate the number of people c_{ij} who commute between any given pair of cities i and j. We use this to

Figure 2. Two travel networks based on empirical data—the US air passenger traffic network, and the Northwest England commuter network. The colour and opacity of the edges scales with their weight, normalized separately in each network relative to the highest weighted edge. The actual weights depend on s. The area of each node scales with the population size it represents, and the colour indicates its betweenness centrality [42,43].

model the travel rate according to the equation:

$$\lambda_{ij} = s \frac{c_{ij} + c_{ji}}{2N_i}, \tag{2.13}$$

where $N_i > 0$ is the population size of city i, and s is an overall scaling factor which we allow to vary. We make the travel rate symmetric because a commuter travelling between cities can transmit a pathogen either from the source city to the destination city (on the outbound journey) or from the destination city to the source city (on the inbound journey) [44]. The key assumption of our models of empirical travel data is that total travel is proportional either to air travel or to commuter travel, respectively.

3. Vaccination strategies

To test different vaccination strategies, we followed three alternative methods of prioritizing regions for vaccination—risk, betweenness centrality and RWPC. We assumed that a pathogen would emerge in a single region at first, putting that region in state I and triggering a vaccination campaign across the network. We performed separate simulations for each possible region of origination of the outbreak, and averaged the results.

We modelled the results of these prioritizations with a range of total available vaccine doses, considering all regions in state S as possible targets for vaccination. We assumed a single-dose vaccine with 100% efficacy against infection. Given a certain number of doses n_d, we assigned enough doses to the highest priority region to achieve herd immunity in that region, then did the same for the second highest priority region, and so on, until we reached a region whose herd immunity threshold was higher than the number of doses remaining, in which case we assigned all of the doses remaining to that region and none to those with lower priority. We assume that the vaccine distribution in a region is always completed before an epidemic occurs in that region.

To determine the priorities of regions according to the risk-targeting strategy, we performed 10 000 stochastic simulations of the outbreak from the initial state, and prioritized regions from most at risk to least at risk. In this study, we treated the risk to region i as equivalent to the frequency with which that region i becomes infected in stochastic simulations. In principle, the risk is the expected number of infected individuals, i.e. the probability of the region becoming infected multiplied by $R(\infty)$. However, both $R(\infty)$ and the number of vaccines required to achieve herd immunity scale linearly with the size of the susceptible population. This

means that the risk averted per dose of vaccine is maximized by prioritizing regions based on the probability of an outbreak in that region, from highest probability to lowest. For the RWPC strategy, we calculated the RWPC for each region according to the initial state of each region, and prioritized regions from highest RWPC to lowest. For the betweenness centrality strategy, we simply prioritized from highest betweenness centrality to lowest betweenness centrality. For both centrality measures, we found that weighting the edges according to the rate of travel per individual resulted in closer to optimal targeting than weighting them according to the total flux of travelling individuals.

4. Results

The results of vaccine distribution on our artificial networks using all three different strategies is shown in figure 3. In each panel, the y-axis shows the mean total number of infections in an epidemic that starts with a single region in state I and the remaining regions in state S, over 10 000 stochastic simulations for each possible starting region, using the model outlined in §2.1. The total number of simulations used to calculate each mean value is therefore 180 000. The smaller the number of infections, the more effective the control of the pathogen. The x-axis in each panel shows the number of vaccines administered, n_d, representing the resources available for control. The three different coloured lines show the expected number of infected individuals when vaccines are distributed according to the risk-targeting strategy (blue), the betweenness centrality strategy (red) or the RWPC strategy (yellow). The panels are arranged left-to-right with increasing values of within-region basic reproduction number R_0, and top-to-bottom with increasing values of adjacent-region travel rate λ. The range of parameters for which simulation results are shown is chosen to illustrate the region around a change in the relative effectiveness of different vaccine strategies, if such a change occurs.

In figure 3a, we see that, for the one-to-one joined grids network, risk-targeting is the best strategy under some parameter combinations and not under others. When both R_0 and λ are small, the risk-targeting strategy is always the most effective of the three considered, for any number of available vaccines. For some combinations of R_0 and λ, such as 8 and 2e−6, respectively, the risk-targeting strategy is only the most effective when there are many vaccine doses available. Finally, when both R_0 and λ are large, the centrality-targeting strategies are more effective than the
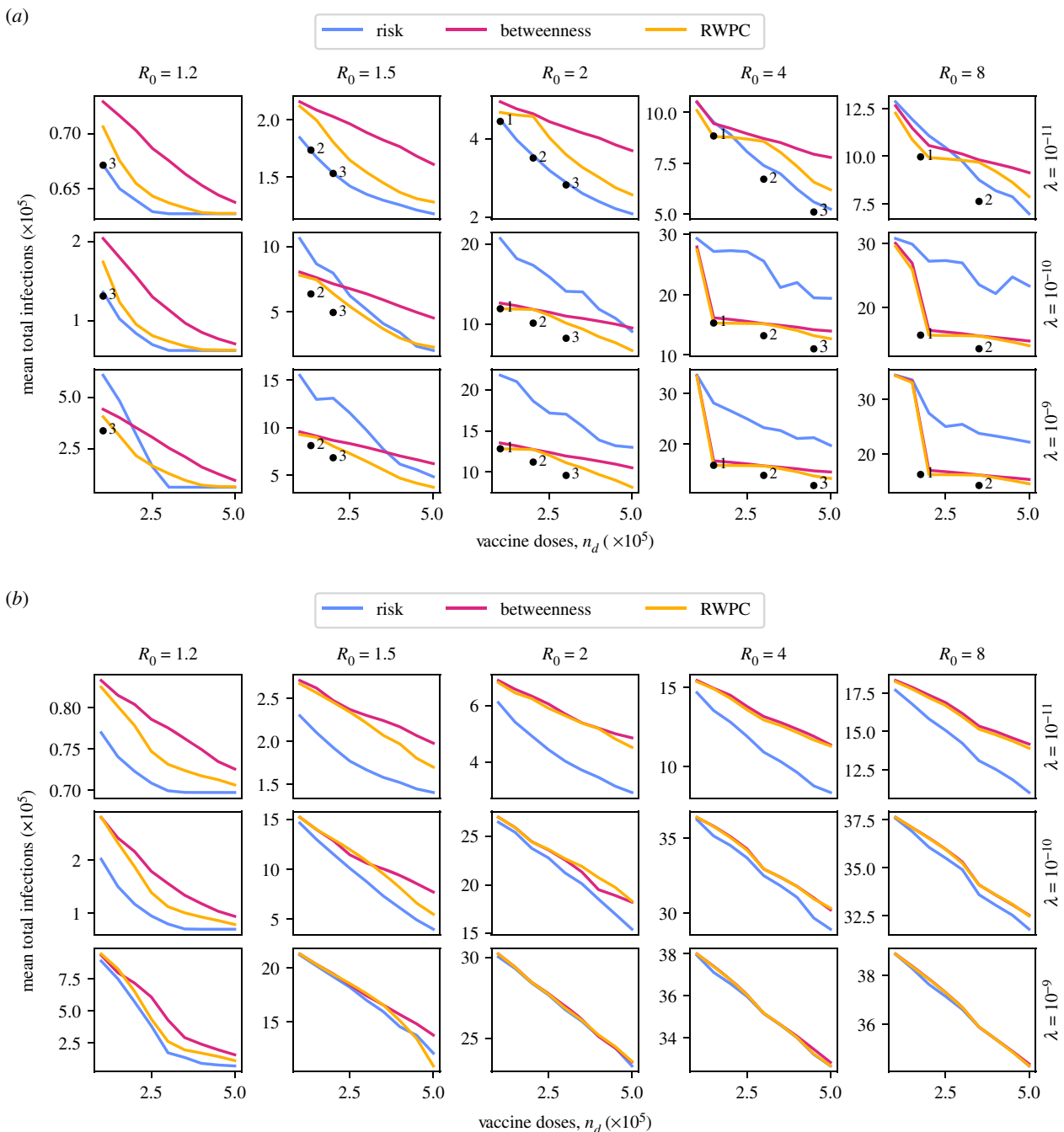
Figure 3. Plots of the mean total number of infections due to an outbreak, against the number of vaccine doses administered according to one of three region prioritization policies, for five different values of R_0 and three different overall travel rates, in two variations on the 'joined grids' network. The long-distance travel parameter a in equation (2.10) is set to 2, and $\mu = 0.5$, in all cases. For the one-to-one variation, it is possible to calculate the optimal vaccine allocation for some parameter values. The resultant mean numbers of infections are shown by black dots, labelled with the number of regions vaccinated to herd immunity at that value of n_d. (a) One-to-one joined grids network, (b) four corners joined grids network.

risk-targeting strategy at all values of n_d. The RWPC strategy is always more effective than the betweenness centrality strategy, with the difference between the two being largest when R_0 and λ are small and n_d is large. The large dips in expected infections seen in the lower right panels of the figure correspond to the point at which sufficient vaccines become available to achieve herd immunity in the patch that joins the two grids. This happens outside of the range of doses considered when $R_0 \leq 2$, and does not have as dramatic an effect when $\lambda = 10^{-11}$. In the electronic supplementary material, we show that increasing a (and thus decreasing long-distance travel) makes risk-targeting more effective, but does not change the general pattern of the results.

In the joined grids network, where each region has the same population size, it is sometimes computationally feasible to determine the optimal allocation of vaccines using a brute force approach. That is, the approach in which we simulate pathogen spread under all possible vaccine allocation prioritizations, and select the prioritization that leads to the smallest expected number of infections. For this approach, we constrain the number of doses to be equal to a small integer multiple of the herd immunity threshold for a single population, so that the number of possible allocations is not excessively high. In this case, we assume that no less than the number of doses required for herd immunity are ever allocated to a region, if any doses are allocated there at
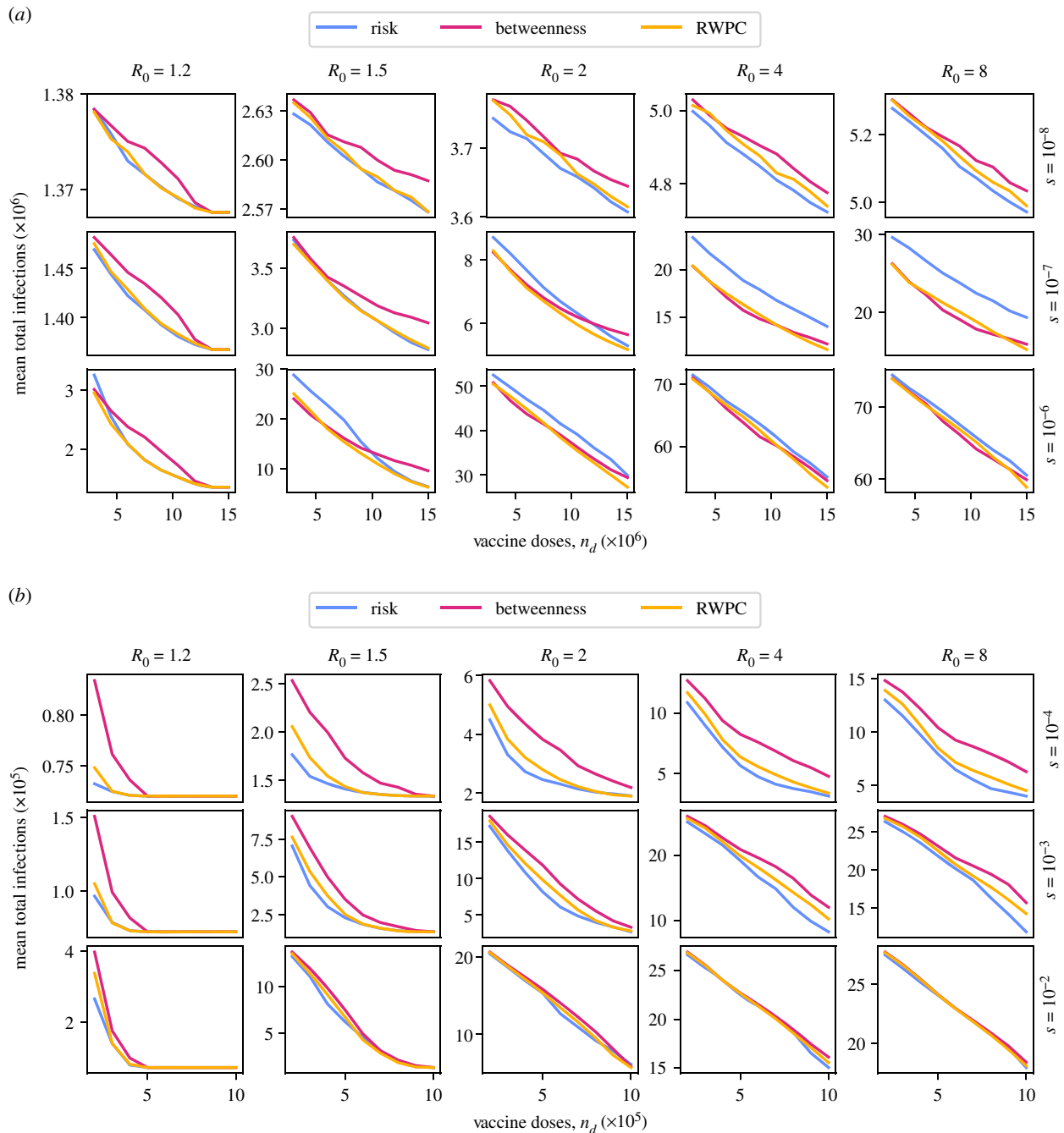
Figure 4. Plots of the mean total number of infections due to an outbreak, against the number of vaccine doses administered according to one of three region prioritization policies, for five different values of R_0 and three different values of the travel scaling factor s, in two different networks based on empirical data. In all cases, we set $\mu = 0.5$. (a) US air passenger traffic network, (b) Northwest England commuter network.

all. The resultant expected number of infections for each of the optimal allocations is indicated with black dots in figure 3a, labelled with the number of regions vaccinated. The spacing between these numbers scales with the herd immunity threshold, as determined by R_0. In general, the best strategy out of those considered for any combination of parameters comes close to optimality in effectiveness, with risk-targeting at low travel rates being indistinguishable from the fully optimal strategy.

In figure 3b, we see that these dynamics can be significantly altered by changing the network structure. When we add multiple routes between the two grids, and increase the population size of two patches, the risk-targeting strategy becomes the most effective strategy for almost all parameter combinations. For very high spread rates the

pathogen infects all those not protected by the vaccine or by regional herd immunity, and so the expected number of infections becomes the same linear function of the number of doses for all vaccine strategies.

In figure 4, we show the results of our various vaccine strategies in empirically derived networks. The structure of the figure is identical to figure 3, but with the scale of travel indicated by the s parameter rather than λ (see equations (2.12) and (2.13) for the role of s in figure 4b and 4a, respectively). The y-axis of each panel again shows the mean number of infections over 10 000 stochastic simulations for each possible starting region. The total number of simulations used to calculate each mean value is therefore 190 000 in the US air passenger traffic network, and 130 000 in the Northwest England commuter network.

In figure 4a, we see that the dynamics of vaccine targeting in the US air passenger traffic network mirrors that in the joined grids network, with the risk-targeting strategy performing best at low R_0 and low inter-regional travel, and the centrality-targeting strategies performing best at high R_0 and high inter-regional travel. The main differences include the overall high performance of the RWPC strategy, which is or is very close to the best strategy in all cases; the convergence of the effectiveness of the various strategies for $s = 10^{-6}$; and the smoothness of curves, with no sharp dips in expected infections. This increased smoothness comes about due to the fact that there is no single bottleneck for pathogen spread in this network, unlike the bottleneck created by the single travel route between the two grids in the one-to-one joined grids network.

In figure 4b, we see a case of an empirical network in which the risk-targeting strategy is consistently the most effective at reducing infections, as seen for an artificial network in figure 3b. The effectiveness of the different strategies converges as s increases.

5. Discussion

In the Introduction, we presented a hypothesis concerning the allocation of vaccine doses to various regions—namely, that risk-targeting would perform better than centrality-targeting when R_0 and inter-region travel rates were high relative to the number of vaccine doses distributed. We find that this hypothesis is borne out by our results, but with the important caveat that only certain network structures allow effective use of centrality-targeting—in networks with many weak edges between highly connected clusters of regions, the effectiveness of centrality-targeting never exceeds that of risk-targeting (figures 3b and 4b).

This provides an important perspective on the widespread use of risk-targeting in spatial vaccination studies. Risk-targeting can perform very well in certain networks, or in certain parameter regimes, but is far from universally superior to alternative strategies. In particular, when a pathogen is spreading quickly, or when vaccine availability is limited, centrality-targeting strategies should be considered as an alternative. However, it is important to choose the correct centrality measure for the system at hand. We have shown that blunt application of a popular measure of node importance, such as betweenness, can result in a less effective allocation of vaccines compared to a measure tailored to the dynamics of a pathogen spreading on a network. In this study, we used a variation on the RWPC measure due to Piraveenan *et al.* [12], but different scenarios may require the application of different measures. For instance, if the current distribution of infections is unknown, percolation centrality cannot be applied and random walk betweenness centrality may be more effective.

The scenarios explored in this study assume a vaccination campaign that can efficiently vaccinate the selected portion of the population before a pathogen becomes widespread. The interacting dynamics of vaccination and transmission for a pathogen that is epidemic or endemic in a population differ significantly from the dynamics explored above [45–47]. Further work is required to know whether risk-targeting is the most effective strategy for combating an endemic pathogen. With SARS-CoV-2 already prevalent in most regions of the world, this means that our results are not applicable to ongoing COVID-19 vaccination efforts; however, they may provide insight into preventing the invasion of new strains of the virus, if these variants are detected early.

In our model, we assumed a single-dose vaccine with 100% efficacy against infection, and no waning immunity. We expect our results to extend to less effective vaccines, since, in our model, reducing the vaccine efficacy is equivalent to reducing the number of vaccinations n_d, so long as herd immunity is still achievable. However, a more complex model of vaccination could expand on our results in interesting ways. An expanded model could consider the different effects of vaccines—infection-reducing, disease-reducing and/or transmission-reducing [48–52]. Varying efficacy in each of these dimensions could affect the results of risk-targeting and centrality-targeting strategies in different ways. Further analyses could also consider different dosing regimes (e.g. vaccines that require multiple doses for maximal efficacy) and the effects of waning vaccine- or infection-induced immunity [53–55].

Vaccination is of course only one way to prevent a pathogen from spreading, and vaccines are not always available early on in an outbreak, especially of a novel pathogen. Risk-targeting is also used when modelling the allocation of prophylactic drugs [9], or the imposition of non-pharmaceutical interventions (NPIs) [56]. In the electronic supplementary material, we extend our framework to a preliminary model of NPIs, and show that our core results are replicated in this case. However, a more complex model is needed to effectively capture the dynamics of different NPIs, and to combine the range of NPIs and pharmaceutical interventions used during emerging outbreaks in a single model. This is a promising avenue for further work.

This study has focused on spatial targeting and has investigated only two types of strategy. In principle, there are as many possible vaccination strategies as there are orderings of individuals to vaccinate. We believe that risk-targeting and centrality-targeting are the best spatial strategies currently in contention, but the vast space of possible strategies is ripe for exploration. As well as variations on centrality targeting, future work could explore generalizations of our results to demographic targeting, investigating when the vaccination of at-risk individuals is most effective relative to those with high spreading capacity (e.g. those with many contacts) [57–59]. Vaccines can also be allocated on a pro rata basis—that is, a fixed proportion of the population is vaccinated in each region. While this strategy has had success in some models, we show in the electronic supplementary material that it is not effective in the disease transmission settings investigated here [2,11].

In this analysis, we have focused on the mean number of infections as the target of control. However, in situations where a large outbreak has an outsized cost, such as when there is some fixed capacity for treatment of a disease, more risk-adverse policies may be more appropriate, in particular those that focus on minimizing some high percentile of the distribution of estimated infections in an outbreak [60,61]. Future work could adapt our model to evaluate risk-adverse vaccination strategies by replacing the exponential distribution of epidemic duration with a more complex probabilistic model of epidemic curves [62]. Such extensions could include a range of pathogen spread dynamics, such as the SEIR or SIRS compartmental models. These different dynamics are likely to impact the temporal variability of the spread of a pathogen, and so are particularly relevant for risk-averse strategies [63].

Finally, analysis of this kind could be applied to realistic models of particular infectious disease outbreaks, providing practical guidance for vaccination campaigns. These applications should consider practical restrictions to vaccine deployment, potentially including penalizing vaccination strategies that have significant spatial complexity. Such work would further elucidate the circumstances in which different types of spatial vaccination targeting strategies would be effective in reducing infections during future outbreaks.

References

1. NHS. Flu vaccine - NHS; 2020. See https://www.nhs.uk/conditions/vaccinations/flu-influenza-vaccine/.

2. Hogan AB et al. 2020 Report 33: Modelling the allocation and impact of a COVID-19 vaccine. Imperial College COVID-19 response team. See https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-33-vaccine/.

3. WHO SAGE Working Group on COVID-19 Vaccines. Roadmap for Prioritizing Population Groups for Vaccines against COVID-19. World Health Organisation. 2020. See https://www.who.int/news-room/q-a-detail/coronavirus-disease-(covid-19)-vaccine-access-and-allocation.

4. Moore S, Hill EM, Dyson L, Tildesley MJ, Keeling MJ. 2021 Modelling optimal vaccination strategy for SARS-CoV-2 in the UK. PLoS Comput. Biol. **17**, e1008849. (doi:10.1371/journal.pcbi.1008849)

5. Kraemer MUG et al. 2017 Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. Lancet Infect. Dis. **17**, 330–338. (doi:10.1016/S1473-3099(16)30513-8)

6. Merler S et al. 2016 Containing Ebola at the source with ring vaccination. PLoS Negl. Trop. Dis. **10**, e0005093. (doi:10.1371/journal.pntd.0005093)

7. Tildesley MJ, Savill NJ, Shaw DJ, Deardon R, Brooks SP, Woolhouse ME, Grenfell BT, Keeling MJ. 2006 Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. Nature **440**, 83–86. (doi:10.1038/nature04324)

8. Keeling MJ, Woolhouse MEJ, May RM, Davies G, Grenfell BT. 2003 Modelling vaccination strategies against foot-and-mouth disease. Nature **421**, 136–142. (doi:10.1038/nature01343)

9. Ferguson NM, Cummings DA, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS. 2005 Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature **437**, 209–214. (doi:10.1038/nature04017)

10. Tuite AR, Fisman DN, Kwong JC, Greer AL. 2010 Optimal pandemic influenza vaccine allocation strategies for the Canadian population. PLoS ONE **5**, e10520. (doi:10.1371/journal.pone.0010520)

11. Venkatramanan S, Chen J, Fadikar A, Gupta S, Higdon D, Lewis B, Marathe M, Mortveit H, Vullikanti A. 2019 Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints. PLoS Comput. Biol. **15**, e1007111. (doi:10.1371/journal.pcbi.1007111)

12. Piraveenan M, Prokopenko M, Hossain L. 2013 Percolation centrality: quantifying graph-theoretic impact of nodes during percolation in networks. PLoS ONE **8**, e53095. (doi:10.1371/journal.pone.0053095)

13. Tanaka G, Urabe C, Aihara K. 2014 Random and targeted interventions for epidemic control in metapopulation models. Sci. Rep. **4**, 5522. (doi:10.1038/srep05522)

14. Holme P. 2017 Three faces of node importance in network epidemiology: exact results for small graphs. Phys. Rev. E **96**, 062305. (doi:10.1103/PhysRevE.96.062305)

15. Kermack WO, McKendrick AG. 1927 A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A **115**, 700–721. (doi:10.1098/rspa.1927.0118)

16. Jagers P. 1975 Branching processes with biological applications. London, UK: Wiley.

17. Keeling MJ, Shattock A. 2012 Optimal but unequitable prophylactic distribution of vaccine. Epidemics **4**, 78–85. (doi:10.1016/j.epidem.2012.03.001)

18. Apolloni A, Poletto C, Ramasco JJ, Jensen P, Colizza V. 2014 Metapopulation epidemic models with heterogeneous mixing and travel behaviour. Theor. Biol. Med. Model. **11**, 3. (doi:10.1186/1742-4682-11-3)

19. Thompson RN, Thompson CP, Pelerman O, Gupta S, Obolski U. 2019 Increased frequency of travel in the presence of cross-immunity may act to decrease the chance of a global pandemic. Phil. Trans. R. Soc. B **374**, 20180274. (doi:10.1098/rstb.2018.0274)

20. Anderson D, Watson R. 1980 On the spread of a disease with gamma distributed latent and infectious periods. Biometrika **67**, 191–198. (doi:10.1093/biomet/67.1.191)

21. Svensson A. 1995 On the asymptotic size and duration of a class of epidemic models. J. Appl. Probab. **32**, 11–24. (doi:10.2307/3214917)

22. Keeling MJ, Rohani P. 2008 Modeling infectious diseases in humans and animals. Princeton, NJ: Princeton University Press.

23. Craft ME, Beyer HL, Haydon DT. 2013 Estimating the probability of a major outbreak from the timing of early cases: an indeterminate problem? PLoS ONE **8**, e57878. (doi:10.1371/journal.pone.0057878)

24. Thompson RN, Gilligan CA, Cunniffe NJ. 2016 Detecting presymptomatic infection is necessary to forecast major epidemics in the earliest stages of infectious disease outbreaks. PLoS Comput. Biol. **12**, 1–18. (doi:10.1371/journal.pcbi.1004836)

25. Gillespie DT. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. **22**, 403–434. (doi:10.1016/0021-9991(76)90041-3)

26. Freeman LC. 1977 A set of measures of centrality based on betweenness. Sociometry **40**, 35–41. (doi:10.2307/3033543)

27. Brandes U. 2001 A faster algorithm for betweenness centrality. J. Math. Sociol. **25**, 163–177. (doi:10.1080/0022250X.2001.9990249)

28. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse HA. 2010 Identification of influential spreaders in complex networks. Nat. Phys. **6**, 888–893. (doi:10.1038/nphys1746)

29. Ahajjam S, Badir H. 2018 Identification of influential spreaders in complex networks using HybridRank algorithm. Sci. Rep. **8**, 11932. (doi:10.1038/s41598-018-30310-2)

30. Christley RM, Pinchbeck GL, Bowers RG, Clancy D, French NP, Bennett R, Turner J. 2005 Infection in social networks: using network analysis to identify high-risk individuals. Am. J. Epidemiol. **162**, 1024–1031. (doi:10.1093/aje/kwi308)

31. Enright J, Kao RR. 2018 Epidemics on dynamic networks. Epidemics **24**, 88–97. (doi:10.1016/j.epidem.2018.04.003)

32. Borgatti SP. 2005 Centrality and network flow. Soc. Netw. **27**, 55–71. (doi:10.1016/j.socnet.2004.11.008)

33. Newman MEJ. 2005 A measure of betweenness centrality based on random walks. Soc. Netw. **27**, 39–54. (doi:10.1016/j.socnet.2004.11.009)

34. Brandes U, Fleischer D. 2005 Centrality measures based on current flow. Lect. Notes Comput. Sci. **3404**, 533–544. (doi:10.1007/978-3-540-31856-9_44)

35. Zipf GK. 1946 The P1P2/D hypothesis: on the intercity movement of persons. Am. Sociol. Rev. **11**, 677–686. (doi:10.2307/2087063)

36. Xia Y, Bjørnstad ON, Grenfell BT. 2004 Measles metapopulation dynamics: a gravity model for

epidemiological coupling and dynamics. *Am. Nat.* **164**, 267–281. (doi:10.1086/422341)

37. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. 2006 Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451. (doi:10.1126/science.1125237)

38. Truscott J, Ferguson NM. 2012 Evaluating the adequacy of gravity models as a description of human mobility for epidemic modelling. *PLoS Comput. Biol.* **8**, e1002699. (doi:10.1371/journal.pcbi.1002699)

39. Bureau of Transport Statistics. Air Carriers: T-100 Domestic Market. 2021.

40. U S Census Bureau. Summary File 1; 2000. See https://www.census.gov/prod/cen2000/doc/sf1.pdf.

41. Office for National Statistics. 2011 Census: Special Workplace Statistics (United Kingdom). 2011.

42. Local_Profil. Blank US Map, Mainland with no States. 2007. See https://commons.wikimedia.org/wiki/File:Blank_US_Map_Mainland_with_no_States.svg.

43. Nilfanion. English ceremonial counties 2010; 2010. Available from: https://commons.wikimedia.org/wiki/File:English_ceremonial_counties_2010.svg.

44. Huang CY, Chin WCB, Wen TH, Fu YH, Tsai YS. 2019 EpiRank: modeling bidirectional disease spread in asymmetric commuting networks. *Sci. Rep.* **9**, 5415. (doi:10.1038/s41598-019-41719-8)

45. Kribs-Zaleta CM, Velasco-Hernández JX. 2000 A simple vaccination model with multiple endemic states. *Math. Biosci.* **164**, 183–201. (doi:10.1016/S0025-5564(00)00003-1)

46. Russell CA, Real LA, Smith DL. 2006 Spatial control of rabies on heterogeneous landscapes. *PLoS ONE* **1**, e27. (doi:10.1371/journal.pone.0000027)

47. Rowthorn RE, Laxminarayan R, Gilligan CA. 2009 Optimal control of epidemics in metapopulations. *J. R. Soc. Interface* **6**, 1135–1144. (doi:10.1098/rsif.2008.0402)

48. Gog JR, Hollingsworth TD. 2021 Epidemic interventions: insights from classic results. *Phil. Trans. R. Soc. B* **376**, 20200263. (doi:10.1098/rstb.2020.0263)

49. Carter R, Mendis KN, Miller LH, Molineaux L, Saul A. 2000 Malaria transmission-blocking vaccines—how can their development be supported? *Nat. Med.* **6**, 241–244. (doi:10.1038/73062)

50. Préziosi MP, Halloran ME. 2003 Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness. *Vaccine* **21**, 1853–1861. (doi:10.1016/S0264-410X(03)00007-0)

51. Smith RJ, Blower SM. 2004 Could disease-modifying HIV vaccines cause population-level perversity? *Lancet Infect. Dis.* **4**, 636–639. (doi:10.1016/S1473-3099(04)01148-X)

52. Harris RJ, Hall JA, Zaidi A, Andrews NJ, Dunbar JK, Dabrera G. 2021 Effect of vaccination on household transmission of SARS-CoV-2 in England. *N. Engl. J. Med.* **385**, 759–760. (doi:10.1056/nejmc2107717)

53. Heffernan JM, Keeling MJ. 2009 Implications of vaccination and waning immunity. *Proc. R. Soc. B* **276**, 2071–2080. (doi:10.1098/rspb.2009.0057)

54. Crellen T *et al.* 2021 Dynamics of SARS-CoV-2 with waning immunity in the UK population. *Phil. Trans. R. Soc. B* **376**, 20200274. (doi:10.1098/rstb.2020.0274)

55. Saad-Roy CM *et al.* 2021 Epidemiological and evolutionary considerations of SARS-CoV-2 vaccine dosing regimes. *Science* **372**, 363–370. (doi:10.1126/science.abg8663originally)

56. Jarvis CI, Gimma A, van Zandvoort K, Wong KL, Edmunds WJ. 2021 The impact of local and national restrictions in response to COVID-19 on social contacts in England: a longitudinal natural experiment. *BMC Med.* **19**, 52. (doi:10.1186/s12916-021-01924-7)

57. Keeling MJ, Rohani P. 2008 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.

58. Wallinga J, Van Bovena M, Lipsitch M. 2010 Optimizing infectious disease interventions during an emerging epidemic. *Proc. Natl Acad. Sci. USA* **107**, 923–928. (doi:10.1073/pnas.0908491107)

59. Gog JR, Hill EM, Danon L, Thompson RN. 2021 Vaccine escape in a heterogeneous population: insights for SARS-CoV-2 from a simple model. *R. Soc. Open Sci.* **8**, 210530. (doi:10.1098/rstb.2020.0263)

60. Tildesley MJ, Smith G, Keeling MJ. 2012 Modeling the spread and control of foot-and-mouth disease in Pennsylvania following its discovery and options for control. *Prev. Vet. Med.* **104**, 224–239. (doi:10.1016/j.prevetmed.2011.11.007)

61. Cunniffe NJ, Cobb RC, Meentemeyer RK, Rizzo DM, Gilligan CA. 2016 Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in California. *Proc. Natl Acad. Sci. USA* **113**, 5640–5645. (doi:10.1073/pnas.1602153113)

62. Thompson RN, Gilligan CA, Cunniffe NJ. 2020 Will an outbreak exceed available resources for control? Estimating the risk from invading pathogens using practical definitions of a severe epidemic. *J. R. Soc. Interface* **17**, 20200690. (doi:10.1098/rsif.2020.0690)

63. Brauer F. 2018 Early estimates of epidemic final sizes. *J. Biol. Dyn.* **13**, 1–8. (doi:10.1080/17513758.2018.1469792)