PNAS RESEARCH ARTICLE | APPLIED MATHEMATICS OPEN ACCESS

# Strength and weakness of disease-induced herd immunity in networks

Takayuki Hiraoka[a,1], Zahra Ghadiri[a], Abbas K. Rizi[a], Mikko Kivelä[a], and Jari Saramäki[a]

Edited by Simon Levin, Princeton University, Princeton, NJ; received October 20, 2024; accepted May 25, 2025

**When a fraction of a population becomes immune to an infectious disease, the population-wide infection risk decreases nonlinearly due to collective protection, known as herd immunity. Some studies based on mean-field models suggest that natural infection in a heterogeneous population may induce herd immunity more efficiently than homogeneous immunization. However, we theoretically show that this is not necessarily the case when the population is modeled as a network instead of using the mean-field approach. We identify two competing mechanisms driving disease-induced herd immunity in networks: the biased distribution of immunity toward socially active individuals enhances herd immunity, while the topological localization of immune individuals weakens it. The effect of localization is stronger in networks embedded in a low-dimensional space, which can make disease-induced immunity less effective than random immunization. Our results highlight the role of networks in shaping herd immunity and call for a careful examination of model predictions that inform public health policies.**

epidemics | herd immunity | network theory

## Significance

During the COVID-19 pandemic, several studies suggested that the spread of infection might induce herd immunity more easily than previously thought due to population heterogeneity. However, these studies relied on differential equation-based epidemic models, which cannot account for correlations between individuals. We reexamine the effect of disease-induced herd immunity using individual-based contact network models. We find that herd immunity is weaker when such correlations are taken into account, so much so that the conclusions of the previous studies may be overturned. This effect is especially pronounced when the contact network is spatially embedded. Our results highlight the importance of considering network effects in policy decisions that affect the lives and well-being of millions in future pandemics.

A key challenge in infectious disease control is to protect a population by conferring immunity (1–6). When some individuals gain immunity and become no longer susceptible to a disease, those without immunity also enjoy a reduced risk of infection because they are less likely to come into contact with others who can transmit the disease. Due to this indirect protection, the presence of immune individuals has a nonlinear impact on the overall level of protection in the population. This concept is known in epidemiology as herd immunity (7–10). In particular, it is often presumed that there is a critical proportion of immune individuals—the herd immunity threshold—above which the chain of transmission cannot be sustained, and hence, the disease cannot invade the population.

Although the term "herd immunity" was coined and a rudimentary understanding of the phenomenon emerged around 1920, it was not until the 1970s that a quantitative theory of herd immunity was developed through mathematical modeling (10–13). The main focus was on estimating the vaccine coverage necessary for disease elimination. A simple expression for estimating the herd immunity threshold was obtained using a basic model that assumes immunization of a homogeneous, fully mixed population (14). At the same time, much theoretical effort has been devoted to bridging the gap between such simplifying assumptions and the heterogeneity of real-world populations (7, 14–19). Such population heterogeneity can be leveraged to design efficient, targeted vaccination strategies.

In general, immunity to an infectious disease can be acquired not only by vaccination but also by previous infection. In their seminal paper (20), Kermack and McKendrick derived what is now known as the final size equation, which implies the notion of disease-induced herd immunity where an epidemic of a disease that confers immunity after recovery can end before the susceptible population is exhausted. More recently, disease-induced herd immunity has gained renewed attention, particularly in the context of the COVID-19 pandemic (21–25), for which no vaccine was available at the early stage. If one assumes a homogeneous and fully mixed population, disease-induced and vaccination-induced herd immunity are mathematically equivalent. However, this equivalence breaks down when variation in contact patterns is introduced. Britton et al. (21) showed that the threshold for disease-induced herd immunity in a heterogeneous population is considerably lower than expected in the homogeneous case; similar conclusions were drawn by several studies that adopted data-driven modeling approaches for COVID-19 (22–25) and more recently for mpox (26, 27). The essential reason for the lower threshold is that individuals with more contacts are more likely to get infected and become immune in an outbreak; epidemic spread thus effectively acts as targeted immunization.

Author affiliations: [a]Department of Computer Science, Aalto University, FI-00076 Aalto, Finland

Author contributions: T.H., Z.G., A.K.R., M.K., and J.S. designed research; T.H. and Z.G. performed research; T.H. and Z.G. analyzed data; and T.H., Z.G., A.K.R., M.K., and J.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: takayuki.hiraoka@aalto.fi.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2421460122/-/DCSupplemental.

Published July 10, 2025.

However, these results are derived using stratified mass-action models defined by differential equations or network models where nodes are randomly linked at every time step. In such modeling approaches, even if population heterogeneities are considered in terms of metapopulations defined by age, household, or spatial separation (28–37), the microscopic structure of persistent interactions between individuals is coarse-grained away and the correlations between the epidemiological states of individuals are disregarded. In other words, these models assume that interactions occur in a mean-field manner where individual details are replaced by averages. While these assumptions provide a convenient starting point, they are seldom met in real-world populations. In reality, interactions often occur repeatedly between the same pairs of individuals and are heavily influenced by social and spatial constraints. Such characteristics are better captured by modeling the contact structure as a static network that encodes these constraints (38–40). In a static network model, the set of individuals one interacts with is assumed to be finite and fixed.

In this study, we use network epidemic models to reexamine how immunity induced by a past epidemic affects the outcome of future epidemics. Building on earlier studies that demonstrate the role of network structure in disease-induced herd immunity (41–47), we aim to unpack the mechanisms that shape herd immunity induced by disease and to highlight the fundamental difference between the network model and mean-field model of herd immunity.

Our key finding is that in network models, disease-induced herd immunity is driven not only by the disproportionate distribution of immunity among socially active individuals—as already identified in mean-field models—but also by another, counteracting mechanism inherent to network models. In an outbreak originating from a single source (an initially infected individual), the set of individuals who become infected and subsequently immune is necessarily topologically contiguous in the network. We refer to this as localization of immunity: Since every immune node after an outbreak is adjacent to at least one other immune node, immunity is strongly correlated between neighbors and topologically clustered in the network compared to randomly distributed immunity; see Fig. 1*A* for a schematic illustration. Consequently, the interface between immune and susceptible nodes is small: there are fewer edges between immune and susceptible nodes than there would be without such localization. This means that susceptible nodes are less protected because fewer of their neighbors are immune, which allows chains of infection that would otherwise be blocked to reach them. They also have more susceptible neighbors to infect, should they become infectious. Such mixing heterogeneities between susceptible and immune subpopulations resemble those discussed in the context of vaccination and other interventions (48–53).

We find that the localization of immunity has a significant impact on herd immunity even in maximally random networks. In homogeneous and/or spatially embedded networks, the effect of immunity localization on herd immunity is even more pronounced and can be strong enough to offset the advantage of disease-induced immunity over randomly distributed immunity. Notably, the localization of immunity cannot be accounted for by mean-field models as typically used in the literature. As a result, mean-field models may overestimate the strength of disease-induced herd immunity.

The rest of the paper is organized as follows. First, we formalize how we quantify the effect of herd immunity in the network setting. We then show that unlike in the mean-field view, in network models disease-induced herd immunity differs from the effect of randomly distributed immunity. In particular, the net effect of disease-induced herd immunity is determined by the competition between two mechanisms: the biased distribution of immunity toward highly connected nodes and the localization of immunity within the network topology. After illustrating the effect of each mechanism in configuration model networks with different levels of heterogeneity in node degree (number of adjacent nodes), we quantify the effect of localization as a function of the level of spatiality of the network. Finally, we illustrate the model specificity of herd immunity by showing that, in epidemic models informed by empirical population structure, a population that the mean-field model considers as completely protected can still be invaded by the disease in the network model.

## Results

In the following, we use herd immunity as a broad term to describe any indirect protection of susceptible individuals against infection by the presence of immune individuals. This includes but is not limited to complete herd immunity associated with a herd immunity threshold, where any sustained transmission of the disease is prevented. Throughout, we assume immunity to be permanent and complete, fully preventing infection and transmission.

We use the canonical susceptible–infected–recovered (SIR) model to describe the dynamics of nonrecurrent epidemics. Individuals and the contacts between them are represented as nodes and edges in an undirected contact network of size $N$. We assume that the contact network is large enough to consider the limit $N \to \infty$. We also assume that the contact network is quenched, i.e., remains static throughout the epidemic timescale. See *Methods* for details on the epidemic model, simulation methods, and analytical calculations.

We are interested in comparing two different scenarios in which immunity is introduced into a fully susceptible population. The first scenario is disease-induced immunity, where individuals gain immunity through contracting the disease. An epidemic spreads from a source node until it eventually dies out, rendering the nodes that have experienced infection permanently and completely immune to future reinfection. In the second scenario, random immunization, susceptible individuals are selected uniformly at random from the network and then permanently and completely immunized. This scenario embodies the outcome of homogeneous vaccination, while it can also be interpreted as the mean-field view of immunity acquired through any means, as there is no correlation between the immunity of neighboring nodes.

Regardless of how immunity is induced, its protective effect at the population level is measured by how much future outbreaks are reduced in size and in likelihood of occurrence. In the case of disease-induced immunity, such post-immunity outbreaks can be caused by a more infectious variant of the original pathogen that induced immunity. Consider now the network of nodes that remain susceptible after the immunity-inducing epidemic or after random immunization. When a fraction $\pi_R$ of nodes are immune, the proportion of this residual subgraph (41–47, 54–59) is $1 - \pi_R$, providing an upper bound for a post-immunity outbreak due to direct protection.

The actual size of the post-immunity epidemic depends on the transmissibility of the disease. However, even if the disease is infectious enough to be always transmitted from an infected individual to a susceptible neighbor, an epidemic caused by a single introduction cannot grow larger than the largest connected component of the residual subgraph. When the size of this largest

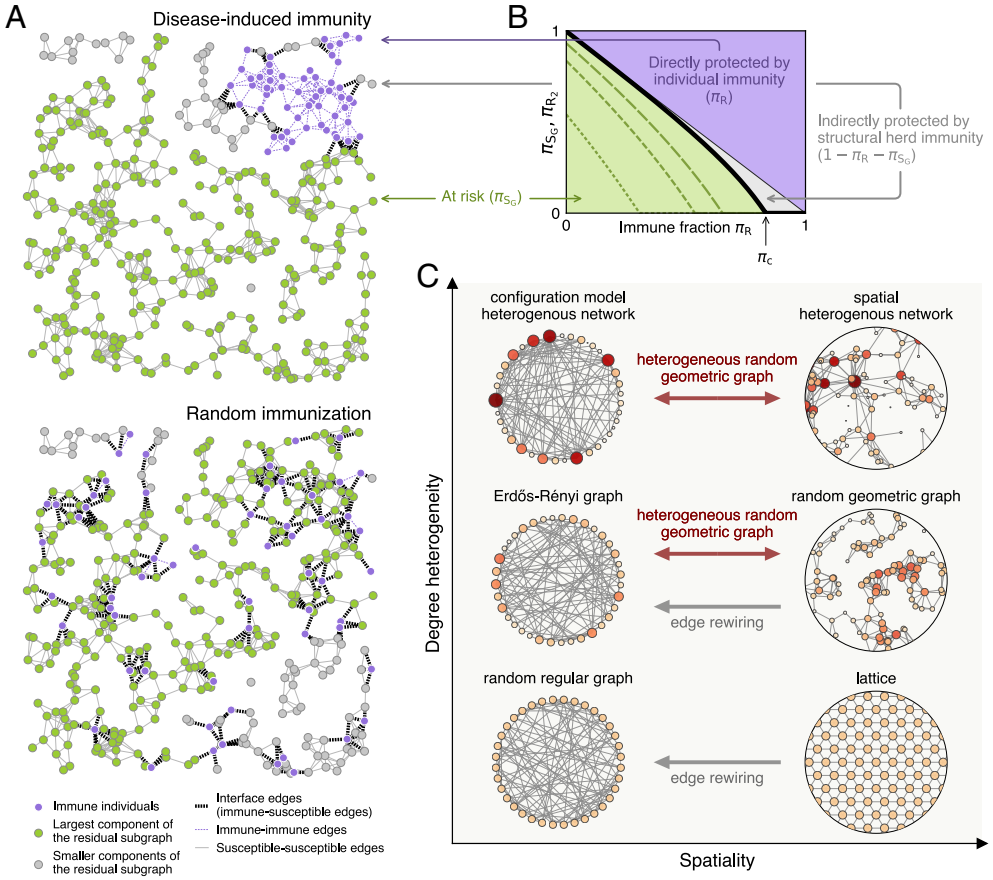


**Fig. 1.** Models. (*A*) Comparison between the distributions of the same number of immune individuals (nodes) induced by disease spreading and by random immunization in a random geometric graph. Immune nodes are colored purple, while susceptible nodes are light green if they belong to the largest component of the residual subgraph and gray otherwise. The edges not included in the residual subgraph are represented by broken lines, with thin purple edges connecting two immune nodes and thick black edges connecting immune nodes and residual nodes. (*B*) Schematic illustration of structural herd immunity characterized by the giant residual component size $\pi_{S_G}$ as a function of the immune fraction $\pi_R$ (thick solid line). The gray area corresponds to the population indirectly protected by structural herd immunity. The green dashed curves are examples of sizes of post-immunity epidemics, with increasing transmissibility from left to right. The intersections of these curves with the horizontal axis indicate thresholds to complete herd immunity; the limiting value $\pi_c$ marks the structural herd immunity threshold. (*C*) Network models used in this study, positioned according to their level of degree heterogeneity (vertical axis) and spatiality (horizontal axis).

connected component scales with $N$, we call it the giant residual component and denote its relative size by $\pi_{S_G}$. Naturally, $\pi_{S_G} \leq 1-\pi_R$, where the equality holds only when the residual subgraph is connected.

Within the class of network models we consider in this work, the giant residual component is the only connected component in the residual subgraph whose size scales with $N$. For large $N$, all smaller connected components in the residual subgraph become negligible in size. Consequently, nodes in these small components are almost certainly shielded from future epidemics of any transmissibility despite not being immune because their risk of exposure is vanishingly small. We refer to this type of indirect protection as structural herd immunity: there are susceptible nodes that are isolated by the immune nodes in the residual contact network.

The effect of structural herd immunity can be quantified by the sum of the sizes of the small components or, equivalently, by the difference between the size of the residual subgraph and its giant component, $1-\pi_R-\pi_{S_G}$ (Fig. 1*B*). If there is no giant component in the residual subgraph (i.e., $\pi_{S_G} \to 0$ when $N \to \infty$), all residual nodes are indirectly protected, and complete herd immunity is achieved regardless of the infectiousness of the disease. The value of $\pi_R$ at which the giant component disappears, denoted by $\pi_c$, signifies the structural herd immunity threshold.

The fraction of nodes $\pi_{S_G}$ in the giant residual component, which represents the maximum possible epidemic size, can vary depending on the specific distribution of immune nodes on the network even if their fraction $\pi_R$ is the same. In particular, we focus on two aspects of how disease-induced immunity is distributed in the network: its bias toward high-degree nodes and its localization. We quantify the first with the ratio $\langle k\rangle_R/\langle k\rangle$ between the mean degree of immune nodes and the mean degree of the entire network. The second aspect, localization of immunity, refers to how adjacent immune nodes are in the network—in a maximally localized configuration, there are as many edges between the immune nodes as possible and in a nonlocalized configuration, immune nodes are located randomly in the network. In the following, we measure the localization of immunity straightforwardly by the share of interface edges between the immune and residual subgraphs, $\rho_{SR} = E_{SR}/E$, where $E_{SR}$ is the number of edges between susceptible and immune nodes and $E$ is the total number of edges in the network. This measure is linearly related to the correlation coefficient between the epidemiological states of adjacent nodes (*SI Appendix*, section A). When immune nodes are more localized, i.e., likely to be adjacent to each other in the network, there are fewer edges at the interface between the two subgraphs. Conversely, in a nonlocalized configuration, the interface is larger

and as the susceptible nodes at the interface have at least one immune neighbor, there are fewer edges to carry a subsequent infection (Fig. 1*A*).

Epidemic dynamics is largely influenced by individual variance in contact and transmission patterns (60–62), which translates to the heterogeneity of node degrees in the contact network. The paradigmatic network model used to express degree heterogeneity is called the configuration model, where the distribution of node degrees solely determines the network structure (63). In a configuration model network, the structure is locally tree-like, meaning that the likelihood of a node being part of a finite-length cycle diminishes as the network size increases. This feature often simplifies analytical calculations and makes the model more tractable.

However, real-world contact networks through which diseases are transmitted are hardly tree-like; rather, they are characterized by the abundance of short cycles. This is because contact and transmission between individuals can only occur when individuals are physically close to each other. If two individuals have a common neighbor in the contact network, they are likely to be near each other, which implies a high probability that they are also connected. As a result, many triangles and short cycles are formed. We will refer to the propensity of individuals to be in contact with other spatially proximate individuals as spatiality.

In this work, we explore a wide range of network structures that differ in terms of degree heterogeneity and spatiality. Degree heterogeneity is characterized by the variance of the degree distribution. For spatial features, we consider that the nodes are endowed with fixed coordinates in a two-dimensional space. Then, spatiality can be measured by the average length of edges (i.e., the average distance between adjacent nodes) in this two-dimensional space. For example, in random geometric graphs (RGGs), edge lengths are short because only local contacts are allowed, representing maximal spatiality, while in Erdős–Rényi random graphs, nodes are in contact independently of their spatial positions, resulting in longer edges on average despite the same Poisson degree distribution. Fig. 1*C* illustrates these two features. See *Methods* for details on the network models used in this study.

**Localization of Disease-Induced Immunity Significantly Weakens Herd Immunity.** We first focus on the structural herd immunity in configuration model networks, where the degree distribution is the only defining feature. We study, in increasing order of degree heterogeneity, random regular graphs (RRGs), Erdős–Rényi random graphs (ERGs), and configuration model networks with negative binomial and power law (scale-free) degree distributions. For each of the network ensembles, we first calculate, both analytically and numerically, the expected size $\pi_R$ of a large epidemic (an outbreak that infects a finite fraction of the population) as a function of transmission rate $\beta$. Then, we compute the size of the giant residual component, $\pi_{S_G}$, for three cases: after removing the nodes that are naturally infected in the epidemic (disease-induced immunity), after removing the same number of nodes but randomly (random immunization), and after removing the same number of nodes with the same degrees but randomly. We refer to the third case as proportional immunization. The analytical calculations are performed using an averaged message-passing framework (41, 64, 65); see *SI Appendix*, section F for details.

The left column of Fig. 2 summarizes our results. For RRGs, we observe that, strikingly, structural herd immunity induced by disease is weaker than that of random immunization; the giant residual component is always larger in the case of disease-induced immunity (Fig. 2*A*). As RRGs have no degree heterogeneity and their structure is entirely random, the strength of herd immunity is entirely dictated by the localization of immunity in the wake of the outbreak. The contiguous nature of the subgraph covered by disease-induced immunity is clearly visible in the smaller size $\rho_{SR}$ of the interface between the immune and residual subgraphs. Conversely, for random immunization, the larger interface is associated with a high level of structural herd immunity, with a residual subgraph whose giant component is smaller.

When the contact network is an ERG where node degrees are moderately heterogeneous, disease-induced immunity and random immunization result in a giant residual component of the same size. Although structural herd immunity is strengthened by the fact that the outbreak disproportionately infects high-degree nodes, it is weakened to an equal extent by the localization of immunity after the outbreak. Both effects are visible in Fig. 2*C*: the average degree $\langle k \rangle_R$ of immune nodes is higher for disease-induced immunity, but the size of the susceptible–immune interface $\rho_{SR}$ is smaller. The equal magnitude of these two effects can be shown analytically; we prove in *SI Appendix*, section G that the Poisson distribution is, in fact, the unique degree distribution of the configuration model in which the two effects exactly offset each other. The giant residual component size under proportional immunization implies that the structural herd immunity would be stronger if only the effect of preference for high-degree nodes were considered: immunizing high-degree nodes wherever they are positioned in the network results in a more splintered residual subgraph. This leads to the important conclusion that if the contribution of localization is neglected, the strength of disease-induced herd immunity is overestimated.

As the degree heterogeneity of the contact network increases, the advantage of disease-induced immunity in exploiting degree heterogeneity and residing disproportionately among high-degree nodes outweighs the localization effect, as seen in the results for configuration model networks with a negative binomial degree distribution with dispersion parameter $r = 3$ (Fig. 2*E*). Importantly, disease-induced herd immunity is still significantly weaker than what would be expected from the degrees of immune nodes alone. This is clearly visible in the difference in the size of the giant residual component for disease-induced immunity and proportional immunization, as well as in the average degrees of immunized nodes and interface sizes. Qualitatively similar results are obtained for scale-free networks, corroborating our findings (Fig. 2*G*). See *SI Appendix*, section K for the robustness of our results against changes in population size and average degree.

To conclude, the net effect of disease-induced herd immunity is determined by the competition between the biased distribution of immunity toward high-degree nodes and the localization of immunity in the network. We observe that localization significantly weakens structural herd immunity in networks with any degree distribution: greater localization of immune nodes leaves the residual subgraph more intact. While the degree heterogeneity of the contact network amplifies the high-degree bias of disease-induced immunity and strengthens herd immunity, it is always counteracted by the effect of localization. Without this countering effect, structural herd immunity would be even stronger, as exemplified by the smaller giant residual component after proportional immunization.

**Spatiality Enhances Localization, Which Further Weakens Disease-Induced Herd Immunity.** Unlike the configuration model networks studied so far, real-world contact networks

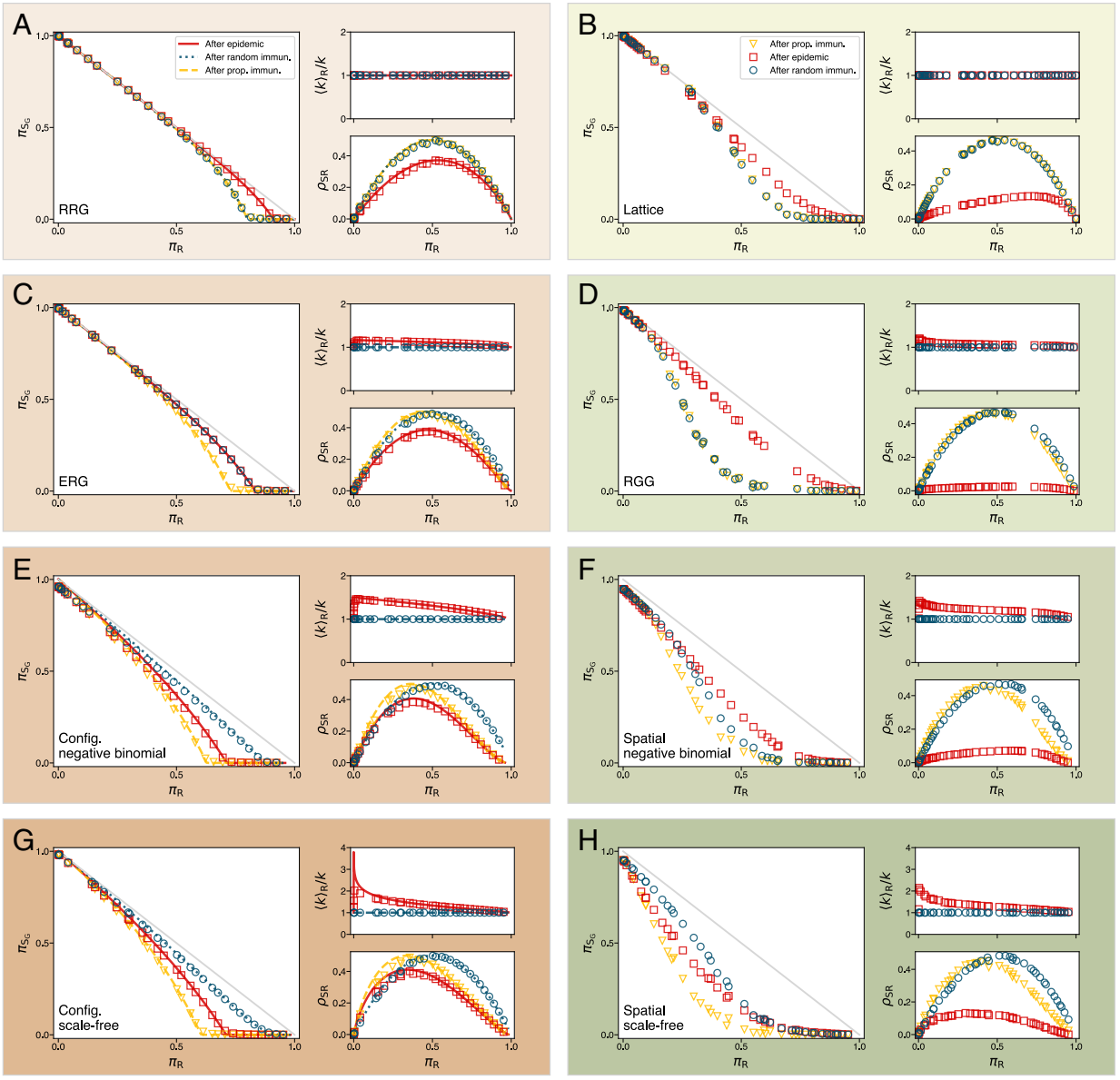


**Fig. 2.** Comparison of the strength of herd immunity for different network models. Each row corresponds to a different degree distribution. Within each row, the left set of panels shows the results for the configuration model network, and the right set of panels shows the results for the corresponding spatial network. (*A*) Random regular graph; (*B*) regular lattice; (*C*) Erdős-Rényi graph; (*D*) random geometric graph; (*E*, *F*) networks with a negative binomial degree distribution (dispersion parameter $r = 3$); (*G*, *H*) networks with a power law degree distribution (exponent $\alpha = 3.1$). For all networks, the number of nodes is $N = 10^4$ and the average degree is $\langle k\rangle = 6$. The heterogeneous spatial networks (*F* and *H*) are generated with a temperature parameter $\tau = 0.05$. For each set of panels, the *Left* panel shows the size $\pi_{S_G}$ of the giant residual component, the *Top Right* panel shows the mean degree of immune nodes normalized by the mean degree of the entire network, $\langle k\rangle_R/\langle k\rangle$, and the *Bottom Right* panel shows the fraction of edges between the immune and residual subgraphs, $\rho_{SR}$. Symbols denote numerical results averaged over 50 different realizations, and lines represent analytical predictions by the averaged message-passing approach. The colors indicate natural infection (red), random immunization (blue), and proportional immunization (yellow). See *Methods* and *SI Appendices*, sections E and F for details of numerical and analytical methods used.

are spatially constrained and are, therefore, effectively low-dimensional. As the ratio of surface area to volume is smaller in lower dimensions, one can expect that the susceptible–immune interface under disease-induced immunity is smaller and that the effect of localization is more pronounced in networks embedded in low-dimensional space. Accordingly, we expect disease-induced herd immunity to be weaker in spatial networks. To study this, we numerically investigate the strength of structural herd immunity in regular lattices, RGGs, and heterogeneous random geometric graphs (HRGGs; see *Methods* for details).

We observe that, for regular lattices, disease-induced immunity leads to a larger giant residual component than random or proportional immunization (Fig. 2*B*), similar to the case for RRGs. The smaller size of the interface, $\rho_{SR}$, implies that the effect of localization of disease-induced immunity is stronger in a regular lattice than in an RRG due to its spatiality.

For RGGs, the gap between disease-induced immunity and random immunization is even larger, confirming the above hypothesis and implying a greater advantage of random immunization over disease-induced immunity in efficiently shrinking the giant residual component (Fig. 2*D*). In an RGG, the immune nodes under disease-induced immunity have a very small interface with the residual nodes, indicating that they are highly localized. This leaves the residual subgraph susceptible to

a future outbreak, as only a few residual nodes have immune neighbors that would break chains of transmission and reduce the size of the largest residual component. Although disease-induced immunity can exploit the modest heterogeneity of the Poisson degree distribution, the impact of localization is much more pronounced, overriding the effect of high-degree bias of immunity. This is in contrast to the case of ERGs, the configuration model counterpart of RGGs, where the effects of the two mechanisms exactly cancel each other out. Note that in RGGs, the two mechanisms are intertwined; high-degree bias amplifies localization because of degree correlations.

The impact of spatiality is particularly evident for networks with higher degree heterogeneity. For a HRGG with a negative binomial degree distribution, structural herd immunity induced by natural infection can be weaker than that achieved by random immunization (Fig. 2*F*). Juxtaposed against the configuration model counterpart, where the opposite result is found, this highlights that the spatiality of the contact network can boost the effect of localization to the extent that it overcomes the counteracting effects of biased distribution of immunity toward high-degree nodes, thus reversing the outcome. In scale-free HRGGs, characterized by an even higher degree heterogeneity, disease-induced immunity still proves more efficient than random immunization in dismantling the giant residual component (Fig. 2*H*). Even then, there is a significant gap between $\pi_{S_G}$ for disease-induced immunity and proportional immunization, highlighting the effect of immunity localization to attenuate structural herd immunity. *SI Appendix*, section K confirms the robustness of our results to changes in population size and average degree.

We have so far established that the spatiality of the contact network diminishes the disease-induced herd immunity by amplifying immunity localization. We next ask: how strongly does the network need to be embedded in space for the effect of high-degree bias of immunity to be outweighed by the effect of immunity localization? To this end, we compare the effectiveness of disease-induced immunity and random immunization across the spatiality spectrum. We use the edge rewiring process and the HRGG model with varying temperatures to cover the spectrum of spatiality continuously. As a measure of spatiality, we use the normalized mean edge length $\langle d\rangle/d^*$ as described in *Methods*. In the configuration model, where the edges are completely random, we have that $\langle d\rangle/d^* = 1$; edges are shorter for networks that are more strongly constrained in space. To estimate the effectiveness of structural herd immunity with a single number, we use the structural herd immunity threshold and measure the minimum fraction of nodes that need to be immune to eliminate the giant residual component: $\pi_c = \min\{\pi_R \mid \pi_{S_G} = 0\}$. In other words, even a disease with an infinitely large transmission rate cannot invade the population if $\pi_R \geq \pi_c$. Thus, $\pi_c$ represents the worst-case bound for the herd immunity threshold, i.e., the structural herd immunity threshold. Here, we numerically identify $\pi_c$ as the minimum value of $\pi_R$ that leads to $\pi_{S_G} \leq 0.01$. We let $\pi_c^{DII}$ and $\pi_c^{RI}$ denote the threshold under disease-induced immunity and random immunization, respectively.

Fig. 3*A* shows the difference between the structural herd immunity thresholds of random immunization and disease-induced immunity, $\pi_c^{RI} - \pi_c^{DII}$, as a function of normalized mean edge length $\langle d\rangle/d^*$ for different degree distributions. Positive values indicate that $\pi_c^{DII}$ is smaller, i.e., disease-induced immunity results in a lower threshold than random immunization, while negative values imply the opposite. For all the degree distributions, the difference between the thresholds increases with spatiality. In other words, the spatiality of the

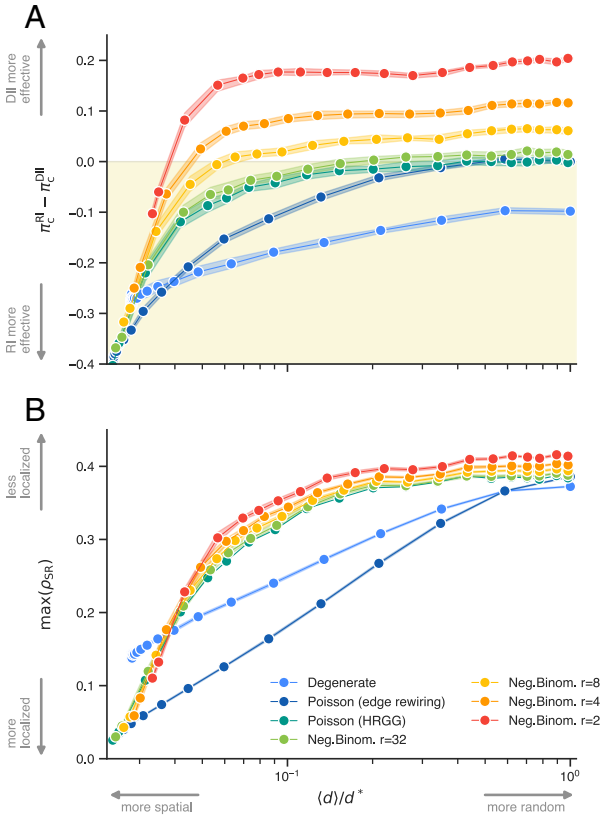


**Fig. 3.** Disease-induced herd immunity as a function of spatiality. Spatiality is measured by the normalized mean edge length $\langle d\rangle/d^*$. (*A*) Difference in the structural herd immunity thresholds $\pi_c^{RI} - \pi_c^{DII}$. Different degree distributions are indicated by line colors. The shaded area along each line represents the 95% CI. Positive values indicate that natural infection induces a stronger herd immunity than random immunization, while negative values suggest the opposite. (*B*) The maximum (peak value) of $\rho_{SR}$ as a function of $\pi_R$ for disease-induced immunity. Smaller values imply stronger localization of immunity. Note that the value of $\pi_R$ at which $\rho_{SR}$ is maximized generally does not coincide with $\pi_c$.

contact network decreases the relative advantage of disease-induced immunity. Fig. 3*B* and *SI Appendix*, section H show that, when the degree distribution is held constant, increasing the spatiality of the network generally shrinks the interface, implying that disease-induced immunity becomes increasingly localized. This increase in localization corresponds with the decrease in the strength of structural herd immunity from natural infection shown in Fig. 3*A*. Therefore, the more spatial the networks are, the more localized disease-induced immunity becomes, and consequently, the weaker the herd immunity acquired through natural infection.

As we have already seen, for Poisson degree distributions, the two immunity scenarios induce an equally strong herd immunity without spatiality (i.e., at $\langle d\rangle/d^* = 1$). For more heterogeneous networks, there is a crossover from the regime where disease-induced immunity is more effective ($\pi_c^{DII} < \pi_c^{RI}$) to the regime where random immunization becomes more advantageous ($\pi_c^{DII} > \pi_c^{RI}$) as the network becomes more spatial. As the degree distribution becomes broader, as in negative binomial distributions with increasing variance (i.e., decreasing dispersion parameter), the crossover point shifts toward the spatial end, indicating that a higher level of spatiality is required to counterbalance the effect of degree heterogeneity (see also *SI Appendix*, section I). This highlights the competition between degree heterogeneity and spatiality of the contact

network in determining the strength of disease-induced herd immunity.

**Comparing Mean-Field and Network Models with Population Structure.** Our analysis above suggests that disease-induced herd immunity in network models is driven by high-degree bias and localization of immunity, while mean-field models only account for the former. This difference between models has real-world consequences for estimating the effect of disease-induced herd immunity. In particular, mean-field models may overestimate the strength of herd immunity and provide an overly optimistic outlook. We demonstrate this by showing, based on empirical contact patterns, that even when the population reaches the condition of complete herd immunity according to a mean-field model, the disease can still reinvade the population in a similarly parameterized network model.

We compare two age-stratified models informed by empirical data: a mean-field model and a network model of SIR dynamics, both with the same age structure and age-specific contact patterns. The age structure is captured by vector $\mathbf{p}$, whose entries are the proportion $p_i$ of the population in each age group $i$. The contact patterns are represented by the contact matrix $M$ whose element $M_{ij}$ represents the average number of contacts an individual in age group $i$ makes with individuals in age group $j$. For the contacts between groups to be symmetric, we impose on the contact matrix $p_i M_{ij} = p_j M_{ji}$.

The mean-field model is defined by the following set of differential equations:

$$\begin{aligned} \dot{s}_i &= -\lambda_i s_i, \\ \dot{y}_i &= \lambda_i s_i - \gamma y_i, \\ \dot{r}_i &= \gamma y_i. \end{aligned} \qquad [1]$$

Here, $s_i, y_i, r_i$ denotes the proportion of susceptible, infected, and recovered individuals in age group $i$, respectively, and satisfies $s_i + y_i + r_i = 1$; $\gamma$ is the rate at which infected individuals recover; $\lambda_i$ is the force of infection to which an individual in age group $i$ is subject, which is calculated as

$$\lambda_i = \beta_{\mathrm{MF}} \sum_j M_{ij} y_j, \qquad [2]$$

where $\beta_{\mathrm{MF}}$ denotes the transmission rate of the mean-field model. In this model, the basic reproduction number can be computed as (28)

$$R_0 = \frac{\beta_{\mathrm{MF}}}{\gamma} \Lambda(M), \qquad [3]$$

where $\Lambda(\cdot)$ denotes the spectral radius. By linearizing Eq. **1** around $s_i$ and $y_i \to 0$, we obtain the condition that disease with transmission rate $\beta_{\mathrm{MF}}$ cannot invade the population:

$$R_0 \frac{\Lambda\big(\mathrm{diag}(\mathbf{s})M\big)}{\Lambda(M)} \leq 1, \qquad [4]$$

where $\mathbf{s}$ denotes the vector of which $i$th element is $s_i$. Thus, this condition represents complete herd immunity of the mean-field model, with the left-hand side representing the effective reproduction number. Note that Eq. **4** is a condition for the vector $\mathbf{s}$ and does not represent a single scalar value of the herd immunity threshold. There are many different $\mathbf{s}$ that make both sides of Eq. **4** equal, and the herd immunity threshold as the total proportion of susceptible individuals in the population, given by $\mathbf{p} \cdot \mathbf{s}$, can vary depending on the specifics of the system and dynamics.

Fig. 4*A* shows a simulated epidemic with a mean-field model of the population of Finland in 2007. The age structure (66) and contact matrix (61, 67) are stratified into 16 age groups. The basic reproduction number is set to $R_0 = 3$ and the mean recovery time is set to $\gamma^{-1} = 5$ d. When the population reaches the herd immunity threshold defined by Eq. **4**, 54.1% of the total population has contracted the disease. However, due to the heterogeneity of contact patterns within the population, there is a large variance in attack rate between age groups, ranging from 10.4% among those over 75 to 73.6% among 25- to 29-y-olds.

We formulate the corresponding network model as follows. The population of $N$ individuals is divided into age groups, each with $n_i \simeq Np_i$ individuals. The contact network is generated by a stochastic block model (SBM). The standard SBM assumes that the probability that individuals are linked to each other depends only on the groups they belong to. It is defined by the number of individuals $n_i$ in each group $i$ and the edge probability $B_{ij}$ between an individual in group $i$ and an individual in group $j$. The edge probability is related to the contact matrix as

$$B_{ij} = aM_{ij}/n_j, \qquad [5]$$

where $a$ is the coefficient that controls the total number of edges in the network. In addition to the age-stratified contact patterns, we further incorporate spatiality by considering a mixture of the SBM and the HRGG model parameterized by temperature $\tau$; see *SI Appendix*, section D for details.

Using the same age structure and contact matrix of Finland, we construct the contact network with $N = 2 \times 10^4$ and $a = 1$. For this setup, the normalized mean edge length $\langle d \rangle / d^*$ decreases nonlinearly with inverse temperature from $\langle d \rangle / d^* = 1$ for $\tau^{-1} = 0$ to $\langle d \rangle / d^* = 0.31$ for $\tau^{-1} = 3$ and $\langle d \rangle / d^* = 0.16$ for $\tau^{-1} = 10$. We let an epidemic of $R_0 = 3$ evolve in the network until it reaches the herd immunity threshold defined by Eq. **4**, at which point we halt the transmission process and let every infected individual recover. Once the population is free of infection, we reintroduce the disease into the population by infecting an individual who remained susceptible during the first epidemic. If the population has already reached herd immunity during the first outbreak, the disease will be quickly eliminated, and there will be no epidemic resurgence that infects a finite fraction of the population. To highlight the maximum risk, we sample post-immunity outbreaks of significant sizes, if they occur at all.

Fig. 4*B* shows the cumulative incidence during the first epidemic and after the reintroduction of the disease according to the network model. The size of the first epidemic is hardly affected by spatiality, varying from an average of 54.2% for $\tau^{-1} = 0$ to 55.2% for $\tau^{-1} = 10$, which is also consistent with the results of the mean-field model. Upon reintroduction of the disease, simulations for standard SBM networks ($\tau^{-1} = 0$) indicate the occurrence of a relatively small epidemic, affecting an additional 2.8% on average. When the network is spatially embedded, we observe a larger resurgence, infecting up to 13.8% of the population. This suggests that the population is still at risk, even though the number of immune individuals should be sufficient to achieve herd immunity according to the mean-field model. In other words, the mean-field model underestimates the herd immunity threshold of networked populations in such a scenario, especially when they are spatially embedded. We discuss the significance of our results for different population sizes in *SI Appendix*, section K and provide an explicit calculation of the network model threshold in *SI Appendix*, section L.

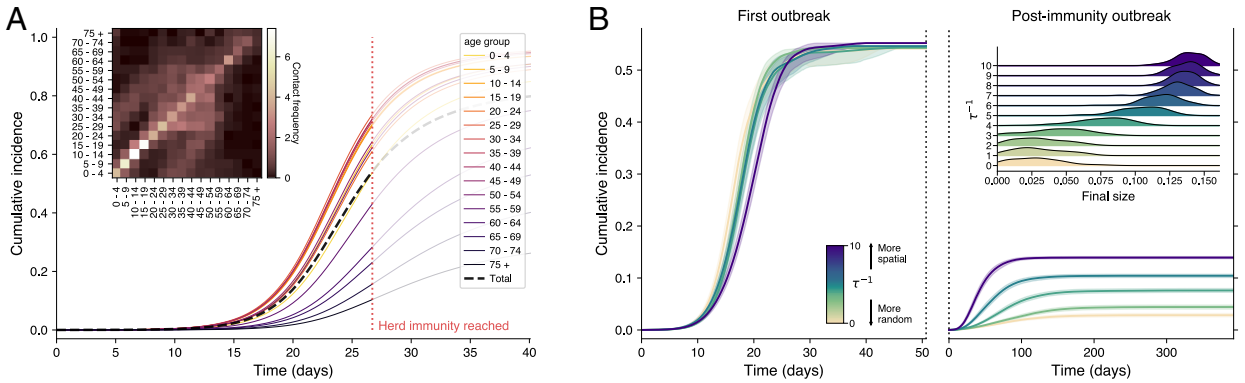


**Fig. 4.** Comparing mean-field and network epidemic models of the empirical population of Finland. (*A*) Mean-field model simulation of an epidemic in an initially susceptible population. The black dashed line represents the cumulative incidence in the entire population, while the solid curves represent the cumulative incidence in each age group, indicated by different colors. The red dotted vertical line indicates the time at which herd immunity is reached. *Inset*: Contact matrix $M$ for Finland's population. (*B*) Cumulative incidence in network model simulations of the first epidemic (*Left*) and after disease reintroduction (*Right*). Solid lines show fixed-time averages over 100 realizations, and shaded areas represent 95% CI. The colors of the curves represent spatial SBM networks with different temperatures: $\tau^{-1} = 0, 3, 4, 5, 10$ from light to dark colors, respectively. The *Inset* in the *Right* panel shows the kernel density estimate of the final sizes of post-immunity outbreaks in networks with different temperatures.

## Discussion

We have studied the effectiveness of disease-induced immunity in network epidemic models. We have found that disease-induced herd immunity is driven by a "tug-of-war" between two antagonistic factors, both stemming from the distribution of immune nodes in the network. On the one hand, an epidemic tends to spread and induce immunity disproportionately among highly connected nodes, which strengthens herd immunity. On the other hand, it leads to a localized distribution of immunity in a topologically contiguous section of the network; this has the opposite effect, that is, it weakens herd immunity, drawing parallels to the effects of inhomogeneous vaccine coverage (51). Importantly, the effect of localization cannot be captured by mean-field epidemic models.

We have shown that the strength of the effects of the two mechanisms (high-degree bias and immunity localization) on disease-induced herd immunity is determined by network structure. In regular networks, where the degrees are identical and only localization is at play, disease-induced immunity provides weaker protection than random immunization. The impact of high-degree bias of immunity increases as the degree distribution becomes more heterogeneous. In Erdős–Rényi graphs (ERGs) that display a moderate amount of degree heterogeneity, the two mechanisms are equally influential, with their effects canceling each other out exactly. In more degree-heterogeneous configuration model networks, disease-induced immunity is more effective than random immunization in conferring protection. However, even in this case, we found that immunity localization attenuates collective protection to a lower level than expected from the degree distribution of immune nodes alone.

When spatiality is added to the picture, the effect of immunity localization is boosted. Notably, even in a degree-heterogeneous network where the distribution of disease-induced immunity is strongly biased toward high-degree nodes, its advantage over random immunization may be offset and overturned because the strong localization weakens the amount of structural herd immunity. We have shown that increasing spatiality generally makes disease-induced herd immunity weaker compared to random immunization. The competition between spatiality and degree heterogeneity of the network is evident from the fact that the former needs to be compensated for by the latter for disease-induced immunity to be more effective than random immunization. In summary, our results suggest that the two underlying mechanisms of disease-induced herd immunity are each influenced by one of the two features of network topology. While degree heterogeneity promotes the effect of the biased distribution of immunity toward high-degree nodes, spatiality enhances the impact of immunity localization.

The difference in the mechanisms of disease-induced herd immunity, namely the presence/absence of the effect of immunity localization, can lead to a discrepancy in the estimated strength of herd immunity between mean-field and network models. We have shown that this inconsistency can arise in stylized models of a real-world population. Even when the mean-field model deems that the population has reached complete herd immunity through natural infection, the disease can reinvade the population in a network model. In other words, the mean-field model estimates a lower threshold and a stronger herd immunity than the network model. Note that this discrepancy is not due to the spatial structure of the network; even when the network model is informed with the same contact data as the mean-field model, its herd immunity threshold is higher.

The tension between the two models stems from the fact that the mean-field model does not account for dynamical correlations between the epidemiological states of individuals. It implicitly assumes that the population is mixed much faster than the epidemic dynamics. The network model, on the other hand, represents the slow-mixing regime; the contact partners of each individual remain the same throughout the epidemic (68). This gives rise to dynamical correlations that manifest themselves in the localized distribution of immune nodes after the spreading of the disease.

In studying network epidemic models, we have based our argument on the notion of structural herd immunity embodied by the residual subgraph and its giant component. Although the notion of giant residual component has been used in previous network literature (41–43), it is not a common practice in epidemiology, where herd immunity is usually characterized against a disease with a specific transmissibility. There are two different ways to interpret what the giant residual component represents. On the one hand, it represents the individuals who will with certainty be infected in a post-immunity epidemic caused by a highly infectious pathogen (i.e., when the transmissibility is

infinite). On the other hand, it represents the population that is still at risk of becoming infected in a post-immunity epidemic of lower transmissibility. When the pathogen is infectious enough to spread in the giant residual component but not infectious enough to infect it entirely, any individual in it can still be infected in the post-immunity epidemic: exactly who will be infected is determined by the stochasticity of seeding and transmission. In contrast, individuals in the smaller components of the residual subgraph are exempt from the risk of infection. This is seen in how the structural constraints affect the final sizes of post-immunity epidemics with specific transmissibility (*SI Appendix*, section J): the results are remarkably consistent with the findings for structural herd immunity.

In a mean-field model, be it a fully mixed mass-action model or a stratified population model parameterized by a contact matrix, an implicit assumption is that every individual interacts with every other individual. This can be mapped to a fully connected contact network (a complete graph). When a subset of nodes is immune and removed from this contact network, the residual subgraph is naturally a complete graph and forms a single connected component. In other words, the residual subgraph is identical to its giant component. This means that i) if the disease is transmissible enough, every individual in the residual subgraph will be infected in a post-immunity epidemic, and ii) some susceptible individuals may not get infected in a post-immunity epidemic of a less infectious disease because of herd immunity, but no one is free from the risk of infection.

We remark that epidemic localization in a network has been discussed in a different context, namely, for the susceptible–infected–susceptible model of recurrent infection (69–72). In these works, the term "localization" implies that infection tends to reside disproportionately on nodes with certain connectivity patterns, such as nodes with high degrees, in inner $k$-cores, and in dense subgraphs. In our work, we focus on nonrecurrent diseases and define localization in terms of strong correlations between the immunity of adjacent nodes as a result of infection spreading over a contiguous region of the network. Despite the different implications of the term, both lines of work recognize the importance of the interplay between network structure and epidemic dynamics.

The concept of disease-induced immunity, particularly its threshold, gained significant attention during the early stages of the COVID-19 pandemic before vaccines became available. Mean-field model predictions suggested that the threshold was lower than expected, leading to policy discussions in some countries on the possibility of protecting high-risk groups by allowing those at lower risk of severe illness to acquire immunity through infection (73–77). In addition to several issues raised by this approach, including ethical concerns, the assumption that at-risk individuals can be identified, and the assumption of permanent and complete immunity (78–80), our analysis suggests that the optimistic projections were, in part, a consequence of the peculiarities inherent in the mean-field models used for estimation.

We emphasize that our aim is not to claim that network models are an accurate depiction of reality. Rather, our results indicate that there is a strong model dependence in the estimation of the strength of herd immunity, which calls for caution in interpreting the results of any mathematical model of infectious disease and warrants careful consideration of its underlying assumptions. The mean-field modeling approach, albeit widely used in epidemic modeling, implicitly assumes a particular set of idealized assumptions that may bias its estimation of disease-induced herd immunity. Outcomes of stylized network models that account for repeated and spatially constrained contacts highlight the need to reassess conclusions drawn from mean-field models in light of more realistic interaction structures. We stress that our network models are also highly simplified and their outcomes are to be taken as theoretical. In this work, we have deliberately disregarded several characteristics of contact patterns in real-world populations in order to single out the effects of two prominent structural features: degree heterogeneity and spatiality.

An advantage of the network approach is that it allows us to build a more individualistic and structural understanding of herd immunity, rather than quantifying it only in terms of thresholds and final epidemic sizes. From this perspective, the role of correlations and inhomogeneities in the network microstructure as well as the presence of mesoscopic structures (81–84) on herd immunity remain to be understood. Furthermore, contacts in real-world populations are not static but associated with specific times, durations, and frequencies. In particular, a large fraction of human interactions occur recurrently with a small set of others (e.g., family, friends, colleagues) while there may be many random encounters of limited duration in public places. In other words, real-world contact networks lie somewhere on the spectrum between the slow and fast mixing regimes. This calls for improved models of contact networks beyond the static, binary ones we have used in our work. Extending our framework for analyzing herd immunity to account for temporality of interactions (85) and heterogeneity in contact frequency and duration is an important goal for future research.

## Methods

**Epidemic Model.** The SIR model on a static network is defined as follows. The dynamical state of each node is either susceptible, infected, or removed, and this state is updated in continuous time. Transmission occurs between each connected pair of an infected node and a susceptible node independently at rate $\beta$, after which the susceptible node becomes infected. Each infected node transitions to the recovered (immune) state at rate $\gamma$. The nodes in the immune state can no longer become infected or transmit the disease and are thus effectively removed from the system. In this work, we set $\gamma = 1$ unless specified otherwise.

In our numerical simulations, we compute the outcome of the SIR model by mapping it to an epidemic percolation network (86), a directed network that exactly encodes the stochastic epidemic dynamics on the original network. This mapping significantly simplifies the numerical analysis; for instance, all the nodes that will be infected in an outbreak can be obtained as the descendants of the initially infected nodes in the epidemic percolation network. The details of the epidemic percolation network framework can be found in *SI Appendix*, section E.

We derive the message-passing formalism for bond percolation (65) to analytically calculate relevant quantities, such as giant residual component $\pi_{S_G}$, average immune node degree $\langle k \rangle_R$, and the boundary size between susceptible and immune nodes, $\rho_{SR}$. For configuration model networks, this method is equivalent to the probability generating function method for solving bond percolation (41, 64). We discuss the details of the message-passing formalism in *SI Appendix*, section F.

**Network Structures.** In this study, we use different network models to control two key topological properties: degree heterogeneity and spatiality. Degree heterogeneity refers to the variance in node connectivity. Node degrees are as homogeneous as possible when all nodes have the same number of neighbors, i.e., the degree distribution is degenerate, as in RRGs and regular lattices. ERGs and RGGs, the degrees are moderately heterogeneous and follow Poisson distributions, where the variance is equal to the mean. At the more heterogeneous end of the spectrum, the network is characterized by the presence of nodes with significantly more connections than average. We use negative

binomial distributions (parameterized by dispersion parameter $r$; a smaller $r$ implies a more heterogeneous degree distribution) and power law distributions (parameterized by exponent $-\alpha$; a smaller $\alpha$ implies a more heterogeneous degree distribution) to represent such high heterogeneity. For all the network models, we fix the mean degree $\langle k \rangle = 6$.

Spatiality refers to the extent to which the underlying geometric configuration of nodes within a low-dimensional metric space determines the connectivity between them. Here, each node occupies a position in the underlying space. In a highly spatial network, the nodes are more likely to be linked to each other if they are closer to each other in space. This is the case for regular lattices and RGGs. At the nonspatial end of the spectrum is the family of configuration network models, such as RRGs and ERGs, representing maximum randomness under the given degree distribution.

We adopt two different approaches to interpolate continuously between spatial and random networks. We use an edge randomization scheme to tune the spatiality for relatively homogeneous networks with degenerate and Poisson degree distributions. Starting with an instance of a network model with the highest spatiality, i.e., a lattice or an RGG, we rewire a fraction $\phi$ of all edges by exchanging the endpoints of two randomly selected edges (87). This process, commonly known as the double edge swap, preserves the degree of each node and allows us to adjust the spatiality without altering the original degree distribution. By completely randomizing edges (i.e., $\phi = 1$), the double edge swap operation transforms a lattice into an RRG and an RGG into an ERG. See *SI Appendix*, section B for details of the edge rewiring process.

To explore the spatial network topologies with higher levels of degree heterogeneity, we adopt the heterogeneous random geometric graph (HRGG) model proposed by Boguñá et al. (88), which allows us to control the degree distribution and spatiality independently. In the HRGG model, each node is assigned an expected degree and a position in a metric space, which is assumed to be a two-dimensional unit square with periodic boundaries. Given the expected degrees and positions, the model generates random networks that satisfy the following properties: i) the degree of each node is a Poisson random variable with expectation equal to the expected degree assigned to the node; ii) the spatiality of the network is expressed as the propensity of nodes to form local connections in the underlying space, which is governed by an independent parameter called the temperature $\tau > 0$. Low values of $\tau$ imply that nodes in proximity are more likely to be linked, and thus, the generated network is more strongly embedded in the space. On the other hand, in the limit of $\tau \to \infty$, the edges are agnostic to the positions of nodes in the underlying space, making the model equivalent to the configuration model. We use this model to generate networks with Poisson and negative binomial degree distributions with varying levels of spatiality. Note that this model cannot generate networks with a degree distribution more homogeneous than Poisson. See *SI Appendix*, section C for details about the HRGG model.

The two methods both induce randomness in the connection patterns between nodes embedded in space, but in different ways. We use the mean spatial length of edges to evaluate the spatiality of the generated networks in a unified manner. The length $d_{ij}$ of the edge between nodes $i$ and $j$ is the Euclidean distance between the positions of $i$ and $j$ in the underlying space, and $\langle d \rangle$ denotes the mean length of all edges. A completely rewired network with $\phi = 1$ and a "hot" HRGG with $\tau \to \infty$ are equivalent to the configuration model; in such cases, the mean edge length is equal to the expected distance $d^*$ between two random points:

$$d^* = 4 \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} \sqrt{x^2 + y^2}\, dxdy = \frac{\sqrt{2} + \ln(\sqrt{2}+1)}{6} \simeq 0.3826\ldots. \quad [6]$$

On the other hand, when $\phi = 0$ or $\tau \to 0$, each node is connected exclusively to other nodes in their proximity; therefore $\langle d \rangle \to 0$ in the large system size limit. Between these two extremes, $\langle d \rangle$ responds monotonically as a function of $\phi$ or $\tau$. We adopt $\langle d \rangle / d^*$ as the normalized measure that represents the randomness of a network with respect to the underlying space.

**Data, Materials, and Software Availability.** Previously published data were used for this work. All the data used for the empirical population simulation are publicly available: the age structure data are available in the Demographic Statistics Database by the United Nations Statistics Division, https://unstats.un.org/unsd/demographic-social/. The contact matrix data are available in the supporting information of Prem et al. (67) with the identifier (https://doi.org/10.1371/journal.pcbi.1005697.s002). The code used in this study is publicly available at https://version.aalto.fi/gitlab/hiraokt1/herd_immunity/.

**ACKNOWLEDGMENTS.** T.H. and J.S. are supported by the Strategic Research Council (SRC) established within the Research Council of Finland (Decision Numbers: 364386 and 364371). M.K. acknowledges support from project 105572 NordicMathCovid, which is part of the Nordic Programme on Health and Welfare funded by NordForsk, and Research Council of Finland project 349366. We wish to acknowledge the Aalto University Science-IT project for generous computational resources.

1. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1991).
2. W. A. Orenstein *et al.*, Measles eradication: Is it in our future?. *Am. J. Public Health* **90**, 1521–1525 (2000).
3. T. A. Reichert *et al.*, The Japanese experience with vaccinating schoolchildren against influenza. *N. Engl. J. Med.* **344**, 889–896 (2001).
4. M. Drolet *et al.*, Population-level impact and herd effects following human papillomavirus vaccination programmes: A systematic review and meta-analysis. *Lancet, Infect. Dis* **15**, 565–580 (2015).
5. J. Toor *et al.*, Lives saved with vaccination for 10 pathogens across 112 countries in a pre-COVID-19 world. *eLife* **10**, e67635 (2021).
6. A. J. Shattock *et al.*, Contribution of vaccination to improved survival and health: Modelling 50 years of the Expanded Programme on Immunization. *Lancet* **403**, 2307–2316 (2024).
7. R. M. Anderson, R. M. May, Vaccination and herd immunity to infectious diseases. *Nature* **318**, 323–329 (1985).
8. R. Anderson, R. May, Immunisation and herd immunity. *Lancet* **335**, 641–645 (1990).
9. R. M. Anderson, The concept of herd immunity and the design of community-based immunization programmes. *Vaccine* **10**, 928–935 (1992).
10. P. Fine, K. Eames, D. L. Heymann, "Herd immunity": A rough guide. *Clin. Infect. Dis.* **52**, 911–916 (2011).
11. P. E. M. Fine, Herd immunity: History, theory, and practice. *Epidemiol. Rev.* **15**, 265–302 (1993).
12. D. Jones, S. Helmreich, A history of herd immunity. *Lancet* **396**, 810–811 (2020).
13. D. M. Morens, G. K. Folkers, A. S. Fauci, The concept of classical herd immunity may not apply to COVID-19. *J. Infect. Dis.* **226**, 195–198 (2022).
14. R. M. Anderson, R. M. May, Directly transmitted infections diseases: Control by vaccination. *Science* **215**, 1053–1060 (1982).
15. J. P. Fox, L. Elveback, W. Scott, L. Gatewood, E. Ackerman, Herd immunity: Basic concept and relevance to public health immunization practices. *Am. J. Epidemiol.* **94**, 179–189 (1971).
16. D. Schenzle, An age-structured model of pre- and post-vaccination measles transmission. *Math. Med. Biol.* **1**, 169–191 (1984).
17. R. M. May, R. M. Anderson, Spatial heterogeneity and the design of immunization programs. *Math. Biosci.* **72**, 83–111 (1984).
18. R. M. Anderson, R. M. May, Spatial, temporal, and genetic heterogeneity in host populations and the design of immunization programmes. *Math. Med. Biol.* **1**, 233–266 (1984).
19. H. W. Hethcote, J. W. Van Ark, Epidemiological models for heterogeneous populations: Proportionate mixing, parameter estimation, and immunization programs. *Math. Biosci.* **84**, 85–118 (1987).
20. W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond., Ser. A* **115**, 700–721 (1927).
21. T. Britton, F. Ball, P. Trapman, A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science* **369**, 846–849 (2020).
22. J. Neipel, J. Bauermann, S. Bo, T. Harmon, F. Jülicher, Power-law population heterogeneity governs epidemic waves. *PLoS ONE* **15**, e0239678 (2020).
23. D. Lu *et al.*, Data-driven estimate of SARS-CoV-2 herd immunity threshold in populations with individual contact pattern variations. medRxiv [Preprint] (2021). https://doi.org/10.1101/2021.03.19.21253974 (Accessed 3 July 2024).
24. M. G. M. Gomes *et al.*, Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *J. Theor. Biol.* **540**, 111063 (2022).
25. R. Aguas, G. Gonçalves, M. U. Ferreira, M. G. M. Gomes, Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. medRxiv [Preprint] (2022). https://doi.org/10.1101/2020.07.23.20160762 (Accessed 3 July 2024).
26. H. Murayama *et al.*, Accumulation of immunity in heavy-tailed sexual contact networks shapes mpox outbreak sizes. *J. Infect. Dis.* **229**, 59–63 (2024).
27. M. Xiridou *et al.*, The fading of the mpox outbreak among men who have sex with men: A mathematical modelling study. *J. Infect. Dis.* **230**, e121–e130 (2024).
28. O. Diekmann, J. Heesterbeek, J. Metz, On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382 (1990).
29. L. Sattenspiel, K. Dietz, A structured epidemic model incorporating geographic mobility among regions. *Math. Biosci.* **128**, 71–91 (1995).

30. A. L. Lloyd, R. M. May, Spatial heterogeneity in epidemic models. *J. Theor. Biol.* **179**, 1–11 (1996).
31. F. Ball, D. Mollison, G. Scalia-Tomba, Epidemics with two levels of mixing. *Ann. Appl. Probab.* **7**, 46–89 (1997).
32. S. Riley, Large-scale spatial-transmission models of infectious disease. *Science* **316**, 1298–1301 (2007).
33. V. Colizza, A. Vespignani, Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* **251**, 450–467 (2008).
34. T. House, M. J. Keeling, Deterministic epidemic models with explicit household structure. *Math. Biosci.* **213**, 29–39 (2008).
35. M. Ajelli *et al*., Comparing large-scale computational approaches to epidemic modeling: Agent-based versus structured metapopulation models. *BMC Infect. Dis.* **10**, 190 (2010).
36. V. Belik, T. Geisel, D. Brockmann, Natural human mobility patterns and spatial spread of infectious diseases. *Phys. Rev. X* **1**, 011001 (2011).
37. C. Zachreson, S. Chang, N. Harding, M. Prokopenko, The effects of local homogeneity assumptions in metapopulation models of infectious disease. *R. Soc. Open Sci.* **9**, 211919 (2022).
38. M. J. Keeling, The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. Lond., Ser. B* **266**, 859–867 (1999).
39. M. J. Keeling, K. T. Eames, Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307 (2005).
40. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
41. M. E. J. Newman, Threshold effects for two pathogens spreading on a network. *Phys. Rev. Lett.* **95**, 108701 (2005).
42. M. J. Ferrari, S. Bansal, L. A. Meyers, O. N. Bjørnstad, Network frailty and the geometry of herd immunity. *Proc. R. Soc. B* **273**, 2743–2748 (2006).
43. S. Bansal, L. A. Meyers, The impact of past epidemics on future disease dynamics. *J. Theor. Biol.* **309**, 176–184 (2012).
44. P. Mann, V. A. Smith, J. B. O. Mitchell, S. Dobson, Two-pathogen model with competition on clustered networks. *Phys. Rev. E* **103**, 062308 (2021).
45. P. Mann, V. A. Smith, J. B. O. Mitchell, S. Dobson, Symbiotic and antagonistic disease dynamics on networks using bond percolation. *Phys. Rev. E* **104**, 024303 (2021).
46. F. Di Lauro, L. Berthouze, M. D. Dorey, J. C. Miller, I. Z. Kiss, The impact of contact structure and mixing on control measures and disease-induced herd immunity in epidemic models: A mean-field model perspective. *Bull. Math. Biol.* **83**, 117 (2021).
47. P. Mann, V. A. Smith, J. B. O. Mitchell, S. Dobson, *N*-strain epidemic model using bond percolation. *Phys. Rev. E* **106**, 014304 (2022).
48. G. Burgio, B. Steinegger, G. Rapisardi, A. Arenas, Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics. *Phys. Rev. Res.* **3**, 033128 (2021).
49. A. K. Rizi, A. Faqeeh, A. Badie-Modiri, M. Kivelä, Epidemic spreading and digital contact tracing: Effects of heterogeneous mixing and quarantine failures. *Phys. Rev. E* **105**, 044313 (2022).
50. G. Burgio, B. Steinegger, A. Arenas, Homophily impacts the success of vaccine roll-outs. *Commun. Phys.* **5**, 70 (2022).
51. T. Hiraoka, A. K. Rizi, M. Kivelä, J. Saramäki, Herd immunity and epidemic size in networks with vaccination homophily. *Phys. Rev. E* **105**, L052301 (2022).
52. H. Watanabe, T. Hasegawa, Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks. *Physica A* **603**, 127760 (2022).
53. A. K. Rizi, R. Michielan, C. Stegehuis, M. Kivelä, Homophily within and across groups. arXiv [Preprint] (2024). https://doi.org/10.48550/arXiv.2412.07901 (Accessed 18 March 2025).
54. S. Bansal, B. Pourbohloul, N. Hupert, B. Grenfell, L. A. Meyers, The shifting demographic landscape of pandemic influenza. *PLoS ONE* **5**, e9360 (2010).
55. S. Funk, V. A. A. Jansen, Interacting epidemics on overlay networks. *Phys. Rev. E* **81**, 036118 (2010).
56. B. Karrer, M. E. J. Newman, Competing epidemics on complex networks. *Phys. Rev. E* **84**, 036106 (2011).
57. T. Hasegawa, N. Masuda, Robustness of networks against propagating attacks under vaccination strategies. *J. Stat. Mech: Theory Exp.* **2011**, P09014 (2011).
58. T. Hasegawa, K. Konno, K. Nemoto, Robustness of correlated networks against propagating attacks. *Eur. Phys. J. B* **85**, 262 (2012).
59. J. C. Miller, Cocirculation of infectious diseases on networks. *Phys. Rev. E* **87**, 060801 (2013).
60. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
61. J. Mossong *et al*., Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).
62. L. Hébert-Dufresne, B. M. Althouse, S. V. Scarpino, A. Allard, Beyond $R_0$: Heterogeneity in secondary infections and probabilistic epidemic forecasting. *J. R. Soc. Interface* **17**, 20200393 (2020).
63. M. Newman, *Networks* (Oxford University Press, 2018).
64. M. E. J. Newman, Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
65. M. E. J. Newman, Message passing methods on complex networks. *Proc. R. Soc. Lond., Ser. A* **479**, 20220774 (2023).
66. United Nations Statistics Division, Population by age, sex and urban/rural residence. *Demographic Statistics Database* (2024). https://unstats.un.org/unsd/demographic-social/. Accessed 15 March 2024.
67. K. Prem, A. R. Cook, M. Jit, Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput. Biol.* **13**, e1005697 (2017).
68. E. Volz, L. A. Meyers, Susceptible-infected-recovered epidemics in dynamic contact networks. *Proc. R. Soc. B* **274**, 2925–2934 (2007).
69. A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, J. F. F. Mendes, Localization and spreading of diseases in complex networks. *Phys. Rev. Lett.* **109**, 128702 (2012).
70. R. Pastor-Satorras, C. Castellano, Eigenvector localization in real networks and its implications for epidemic spreading. *J. Stat. Phys.* **173**, 1110–1123 (2018).
71. G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, L. Hébert-Dufresne, Master equation analysis of mesoscopic localization in contagion dynamics on higher-order networks. *Phys. Rev. E* **103**, 032301 (2021).
72. G. St-Onge *et al*., Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs. *Commun. Phys.* **5**, 25 (2022).
73. P. G. T. Walker *et al*., The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* **369**, 413–422 (2020).
74. E. J. W. Orlowski, D. J. A. Goldsmith, Four months into the COVID-19 pandemic, Sweden's prized herd immunity is nowhere in sight. *J. R. Soc. Med.* **113**, 292–298 (2020).
75. N. Perra, Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Phys. Rep.* **913**, 1–52 (2021).
76. N. Brusselaers *et al*., Evaluation of science advice during the COVID-19 pandemic in Sweden. *Hum. Soc. Sci. Commun.* **9**, 91 (2022).
77. M. Zenone, J. Snyder, A. Marcon, T. Caulfield, Analyzing natural herd immunity media discourse in the United Kingdom and the United States. *PLoS Glob. Public Health* **2**, e0000078 (2022).
78. T. S. Brett, P. Rohani, Transmission dynamics reveal the impracticality of COVID-19 herd immunity strategies. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 25897–25903 (2020).
79. L. F. Buss *et al*., Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science* **371**, 288–292 (2021).
80. D. Sridhar, D. Gurdasani, Herd immunity by infection is not an option. *Science* **371**, 230–231 (2021).
81. C. Stegehuis, R. Van Der Hofstad, J. S. H. Van Leeuwaarden, Epidemic spreading on complex networks with community structures. *Sci. Rep.* **6**, 29748 (2016).
82. L. Hébert-Dufresne, A. Allard, Smeared phase transitions in percolation on real complex networks. *Phys. Rev. Res.* **1**, 013009 (2019).
83. S. Morita, Solvable epidemic model on degree-correlated networks. *Physica A* **563**, 125419 (2021).
84. F. Ball, L. Critcher, P. Neal, D. Sirl, The impact of household structure on disease-induced herd immunity. *J. Math. Biol.* **87**, 83 (2023).
85. P. Holme, J. Saramäki, Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
86. E. Kenah, J. M. Robins, Second look at the spread of epidemics on networks. *Phys. Rev. E* **76**, 036113 (2007).
87. B. K. Fosdick, D. B. Larremore, J. Nishimura, J. Ugander, Configuring random graph models with fixed degree sequences. *SIAM Rev.* **60**, 315–355 (2018).
88. M. Boguñá, D. Krioukov, P. Almagro, M. Á. Serrano, Small worlds and clustering in spatial networks. *Phys. Rev. Res.* **2**, 023040 (2020).