Population structure across scales facilitates coexistence and spatial heterogeneity of antibiotic-resistant infections

Madison S. Krieger[1,*], Martin A. Nowak[1], and Alison L. Hill[1]

[1]Program for Evolutionary Dynamics, Harvard University, Cambridge MA 02138
[*]mkrieger@fas.harvard.edu

ABSTRACT

Antibiotic-resistant infections are a growing threat to human health, but basic features of the eco-evolutionary dynamics remain unexplained. Most prominently, there is no clear mechanism for the long-term coexistence of both drug-sensitive and resistant strains at intermediate levels, a ubiquitous patterns seen in surveillance data. Here we show that accounting for structured or spatially-heterogeneous host populations and variability in antibiotic consumption leads to persistent coexistence, which is robust over a wide range of treatment coverage, drug efficacies, costs of resistance, and mixing patterns. Moreover, this mechanism can explain other puzzling spatiotemporal features of drug-resistance epidemiology that have received less attention, such as large differences in the prevalence of resistance between geographical regions with similar antibiotic consumption or that neighbor one another. Our analysis identifies key features of host population structure that can be used to assess resistance risk and highlights the need to include spatial or demographic heterogeneity in models to guide resistance management.

Introduction

Antibiotic resistance is a major threat to our ability to treat bacterial infections. Over the past century, resistance to each new class of drugs has appeared soon after clinical use began. Today, drug-resistant infections are estimated to cost perhaps $20 billion annually[1]. Individual bacteria that are resistant to multiple classes of antibiotics are now common in species such as *Streptococcus pneumoniae*, *Pseudomonas aeruginosa*, and *Clostridium difficile*[2], and nearly untreatable strains of *Neisseria gonorrheae* [3], *Klebsiella pneumoniae* [4], and *Acinetobacter baumannii*[5] have recently been identified. These trends have led to speculations about a "post-antibiotic future", in which routine medical procedures such as surgeries, childbirth, and dental work might become as high-risk as they were in the pre-WWII era due to a lack of effective antibiotic prophylaxis and treatment[6,7].

Beyond sensational headlines, there is widespread interest among experts in predicting the future morbidity, mortality, and economic impact of drug-resistant infections, so that appropriate investments to counter these trends can be encouraged from government and industry. Mathematical modeling has traditionally played a key role in predicting the dynamics of infectious diseases[8], and has a long history of application to antibiotic-resistant infections[9]. Despite this, there are several recurrent trends in the spatiotemporal dynamics of drug-resistant bacteria that are difficult to explain. Firstly, for many bug-drug pairs, resistant strains have not completely displaced drug-sensitive ones, but instead coexist stably at intermediate levels, in contrast to predictions of standard infection dynamics and the ecological principle of competitive exclusion (Figure 1i). Secondly, regions which prescribe similar levels of antibiotic can have very different levels of resistance, and these differences persist over time (Fig. 1ii). Finally, even neighboring regions can have large and persistent differences in resistance frequencies, at the scale of neighborhoods all the way up to countries (Fig. 1iii). Since models must, at minimum, be able to explain current trends before being trusted for forecasting, these fundamental disagreements with data have either discouraged efforts to make predictions or led to widespread suspicion of existing predictions (e.g. the Review on Antimicrobial Resistance[10,11]).

In this paper we propose a general mechanism to explain these perplexing spatiotemporal trends in antibiotic resistance levels that have eluded previous mathematical models. We consider that host populations may be structured, with heterogeneities in mixing patterns within and between regions, as well as in the distribution of antibiotic use. Spatial differences in treatment rates could arise, for example, from local differences in prescribing guidelines or norms, availability of care, or care-seeking behavior; from the presence of hospitals or other facilities with higher rates of antibiotic usage, or, an increased use of antibiotics following spatially-localized viral epidemics. We find that in contrast to well-mixed populations, structured populations often support long-term coexistence of drug-resistant strains at intermediate levels. We observe that the predicted prevalence of resistance depends not only on the frequency of antibiotic use, but also on how drug use is distributed and details of the transmission network. This mechanism also reproduces the observation of sharp gradients in resistance levels between neighboring regions, even when infections can spread directly between them. We discuss how these ideas can be used to better understand factors

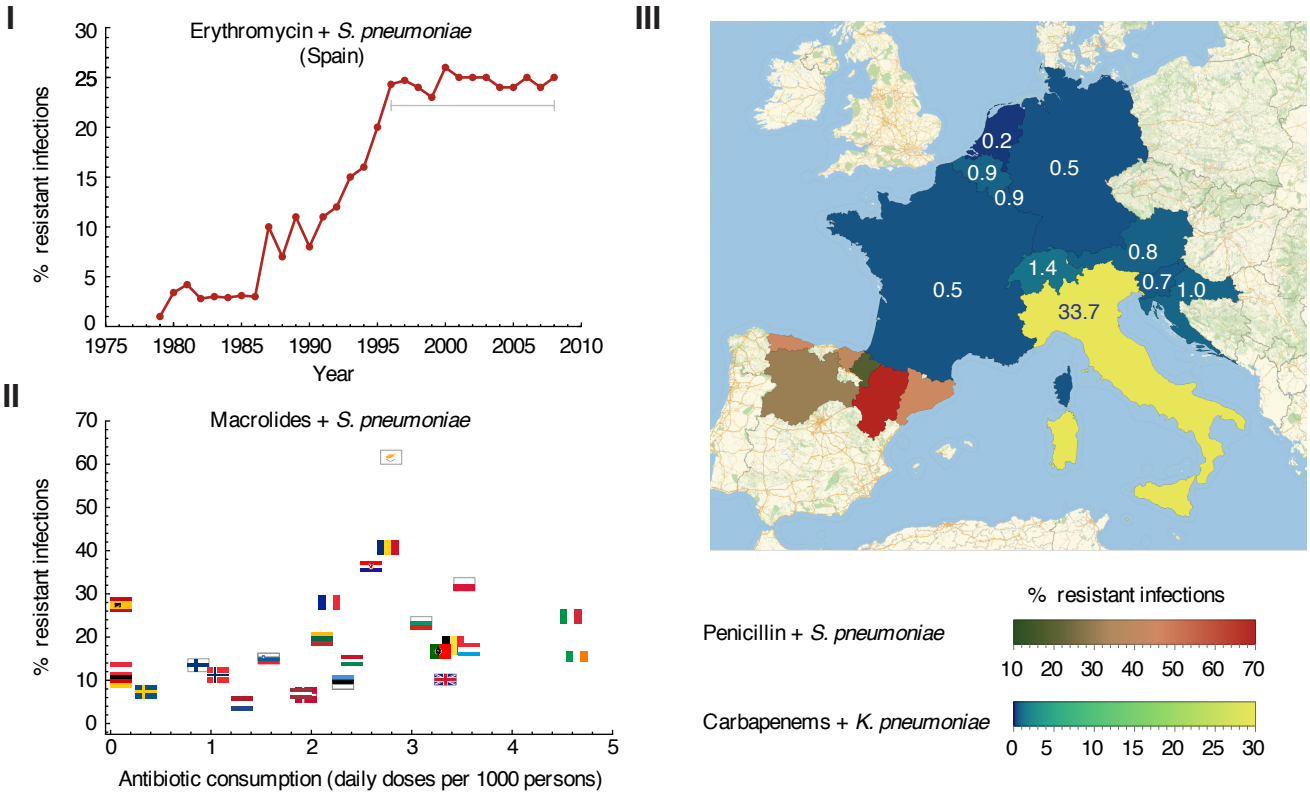

Figure 1. A selection of data illustrating spatiotemporal patterns of antibiotic resistance that have eluded mathematical models. I) Percentage of *S. pneumoniae* isolates in Spain which are resistant to erythromycin (a macrolide), over time. After an initial increase, drug-sensitive and drug-resistant strains have appeared to coexist stably for $\sim$ 15 years.[12]. **II)** The same amount of antibiotic use can lead to very different levels of resistance in different regions. Percentage of *S. pneumoniae* isolates from 2005 that were resistant to penicillin versus the number of daily doses of penicillin administered per capita, for each country participating in the European Centres for Disease Prevention and Control[13,14]. **III)** Neighboring regions can have vastly different rates of resistance which persist for long time periods. Time-averaged percent of *K. pneumoniae* isolates resistant to carbapenems in Austria, Belgium, Croatia, France, Germany, Italy, Luxembourg, and Slovenia from 2013-2016 and Switzerland from 2015-2016[13,15] (blue/yellow), along with *S. pneumoniae* isolates resistant to penicillin in three provinces of Spain from 1990-1998 (green/red)[16,17]. The year-to-year deviation from this average is less than 3% (see Fig. S1). Each country is labeled with the resistance level (%). Note that for *S. pneumoniae* and macrolides or penicillin, "% resistant" includes isolates classified as "non-susceptible".

promoting or hindering antibiotic resistance and to construct more realistic models to evaluate antibiotic stewardship policies.

Observed spatiotemporal trends in antibiotic resistance

I) Frequencies of resistance over time seem to defy competitive exclusion

For a very large number of bacteria-antibiotic combinations, the prevalence of resistant strains has not reached 100% in the population, but has rather appeared to saturate at an intermediate frequency (Figure 1i). Prominent examples include *Escherichia coli* and aminopenicillins in Europe over the past decade[13,18], methicillin-resistant *Staphylococcus aureus* in the United States[2], and penicillin resistance in *S. pneumoniae* in Spain from the 1980's to 1990's[16,17] as well as in other locations/timeframes[19,20]. This long-term coexistence of drug-sensitive and drug-resistant strains is difficult to explain. Resistant strains are clearly selected for in treated individuals, but generally carry a fitness cost such that sensitive strains do better in untreated individuals[21–23]. Mathematical models of infection dynamics under treatment predict that in most conditions the population will move towards an equilibrium where only a single strain persists, with drug-resistant strains reaching fixation when treatment is common enough and costs of resistance are low enough, and drug-sensitive strains dominating otherwise[24]. More generally, long-term coexistence of two strains would seem to defy the ecological principle of competitive exclusion[25–27], which dictates that when two species compete for a single resource (here, the susceptible population), only one may survive.

Dozens of previous mathematical models have been constructed in an attempt to explain this coexistence. While many studies have focused in particular on *S. pneumoniae*, these and others have attempted to be as general as possible so the conclusions can be extended to other bacterial species. Studies have repeatedly found that well-mixed populations in which individuals can only be infected with one strain (either sensitive or resistant) at a time do not generally support coexistence[24,28–30]. Models that include more details of within-host infection dynamics have suggested some potential mechanisms of coexistence. Drug resistant and sensitive strains can coexist if dual infection of individuals with both is possible, but only for a very narrow range of values for both the treatment coverage level and the cost of resistance[24,28,29]. If *de novo* generation of one strain in an individual infected with the other can occur (presumably via point mutation or horizontal gene transfer from non-pathogenic co-colonizers), then coexistence is again possible[24], but such a process is likely to lead to low-level mutation-selection balance, not coexistence at near intermediate levels. Recently, Davies et al.[30] have suggested that co-colonization of hosts with subsequent within-host competition results in a type of frequency-dependent selection that helps maintain coexistence. This occurs because resistant strains have an advantage when they are rare, because they will mostly co-colonize drug-sensitive hosts who, when treated, will be rid of competitors. This mechanism is quite robust to parameter values, but only relevant to commensal bacteria for which co-colonization is common.

Another set of mechanisms shown to promote coexistence is host population structure. If there are two completely isolated sub-populations[24], or if treated and untreated individuals interact extremely rarely[28], then sensitive and resistant strains can both persist, but this is unlikely to apply to any realistic scenario. Simulating transmission in a population in which both contact patterns and the probability of being treated are age-dependent and informed by data for *S. pneumoniae* reveals slightly more opportunity for coexistence[29]. However, this explanation is still far from general, since coexistence only occurs for costs of resistance $<8\%$, does not display full the range of resistance prevalence levels — and their dependence on antibiotic use — observed in data, and is somewhat reliant on the division of the bacterial population into discrete serotypes which each support different resistance levels.

There are also ecological effects known to promote coexistence of species more generally[31], but their relevance to bacterial resistance to antibiotics remains unclear. For example, coexistence can occur when species compete more strongly among themselves than with other species[28]. The obvious candidate for such a mechanism in the context of infectious diseases is strain-specific immunity[32], with the effect that hosts are less susceptible to re-infection by a strain with which they have previously been infected. Strain-specific immunity leads to balancing selection, since low frequency strains have an advantage. However, there is generally not expected to be a connection between the resistance status of a strain and its immunogenicity (e.g. serotype). Lehtinen et al.[33] have recently shown that such a connection may not be necessary, as linkage between a locus under balancing selection and a polymorphic locus that influences the relative fitness of resistant and sensitive strains (such as duration of carriage for *S. pneumoniae*) can promote coexistence. The relevance of these mechanisms to most antibiotic resistant bacteria remains uncertain.

Overall, despite some progress in explaining the long-term coexistence of drug-sensitive and drug-resistant strains, many proposed mechanisms only reproduce coexistence for small regions of parameter space or species-specific scenarios, and the ultimate set of causes for real-world coexistence is far from fully understood.

II) Different frequencies of resistance are seen in regions with similar rates of drug prescription

Another puzzling trend in spatiotemporal antibiotic resistance data is the appearance of very different levels of resistance in regions that seem to have the same levels of drug usage. While overall, there is a strong correlation between the amount of antibiotics consumed in a region and the prevalence of drug resistant strains[14,34], there are many cases that diverge from this trend. For instance, comparing countries participating in the European Centre for Disease Prevention and Control (ECDC), the rate of resistance can vary by as much as 30% even between countries that have equivalent rates of outpatient prescription of penicillin, such as Spain and Poland[13,14] (Figure 1ii). Other examples can be found at smaller spatial scales, for instance between cities in the American South[35] and provinces in Spain[17].

While this finding has not yet been specifically examined in the context of theoretical models, all of the models mentioned in the previous section[24,28–30,33] admit only one stable equilibrium for a given set of parameters, including antibiotic consumption level. Therefore, if each region is considered to be an isolated and well-mixed population, extra mechanisms are needed to explain different resistance levels for the same overall antibiotic use. It could be that one or both of the regions in question had not yet reached a steady-state prevalence of resistance by the time of sampling, though when longitudinal surveys of these same resistant strains are available, they generally suggest that levels have approximately equilibrated (e.g. ECDC data[13]). Another possibility is that the cost of resistance varies between regions, for example due to different mechanisms of resistance or different genetic backgrounds, but experimental evidence to support or refute this idea is lacking so far. Finally, the discrete regions considered during resistance surveillance are in reality neither completely isolated nor well-mixed. Movement of individuals and microbes between regions leads to interdependence of resistance dynamics. Within a region, there could be heterogeneous distributions of treatment and infections which impacts the overall resistance level. Consequently, models that

attempt to explain this trend must consider the connected nature of regions across spatial scales.

III) Neighboring regions can show very different frequencies of resistance

Another confounding element of antibiotic resistance epidemiology is the high degree of heterogeneity in resistance levels between neighboring populations at many differing scales. This is slightly different from the preceding point (II): not only do we find two regions with the same levels of antibiotic use but with different levels of resistance, we find that these regions can be bordering each other. Neighboring regions are likely to exchange infected individuals frequently, which would be expected to ameliorate differences in resistance levels over time, just like the predictions of well-mixed models. In contrast, they can actually sustain sharp gradients in the frequency of resistance seemingly indefinitely.

The scales at which this phenomenon is observed range from the size of nations down to neighborhoods of a city (Figure 1iii). For instance, in Europe, the frequency of carbapenems resistance in *K. pneumoniae* has been much much higher in Italy than in neighboring states for many years[13]. Similar trends have led to differences in the frequency of certain resistant strains across the United States[36,37]. At a smaller scale, in the long-term study of *S. pneumoniae* strains in Spain revealed a large difference in frequency between penicillin resistance in Aragon and its neighboring regions[12,16,17,38]. At an even smaller scale, recent (2015) data on the United States revealed that in four neighborhoods in Greater Miami, rates of cefazolin resistance in *Proteus mirabilis* differed by a staggering 89% overall (rates of 8%, 13%, 88%, 97%)[35]. The fact that these large differences in resistance levels appear frequently in data at different spatial scales and between neighbors that disburse similar amounts of drug therefore require a unifying explanation.

Results

A general model for the evolution of drug resistance in a structured population

To better understand the mechanisms that could be responsible for the spatiotemporal motifs seen in drug resistance data, we developed a simple model for competition between strains of an infection in a structured population. We assume that the total population is divided up into subpopulations of uniform size ('demes'[39], also known as 'households'[40]) (Figure 2). These demes could represent any subdivision of a human population of interest, such as different countries, regions within a nation, neighborhoods within a city, demographic groups, and so forth. Within this population we consider the concurrent spread of drug-sensitive and drug-resistant strains of an infection. Infected individuals spread the disease to uninfected individuals at rate κ if they are in the same deme ('within-deme' transmission rate) and rate β if they are in different but connected demes ('between-deme' transmission rate), with $\kappa \geq \beta$. We do not allow for any super-infection or co-infection: only susceptible individuals can be infected. Infected individuals recover and become susceptible again at rate g.

Upon infection, drug treatment may be administered. For individuals infected with the drug-sensitive strain, treatment reduces the rate at which they transmit the infection to others (in the same or connected demes) by a factor $(1-\varepsilon)$, where ε is the efficacy of the drug, $0 < \varepsilon < 1$. We assume that the resistant strain is perfectly resistant to the drug, so that transmission rates of individuals infected with it are unaffected by the presence of treatment. However, the drug-resistant strain incurs a fitness cost c, which results in lower transmission rates relative to the wild-type strain with or without treatment: $\kappa \to \kappa(1-c)$, $\beta \to \beta(1-c)$. Another possibility for the effect of treatment is that it accelerates disease recovery, and we consider this possibility in the **Supporting Information**.

Our model is an extension of the classic Susceptible-Infected Susceptible (SIS) model for disease transmission[41] to a case with two disease strains (sometimes called an SI_1I_2S model[42]). We assume that each deme is large enough that the dynamics can be considered deterministically. This model is structurally neutral[43], meaning that if two identical strains are simulated (e.g. if $\varepsilon = 0$ and $c = 0$), they will continue to persist *ad infinitum* at the same level at which they were initiated.

Absence of coexistence in an unstructured population

We first consider the dynamics of competition between drug-resistant and drug-sensitive strains in a single large, well-mixed deme (an unstructured population) (Figure 3A). A fraction f of all individuals are treated with drug immediately upon infection by either strain. The dynamics are described by a set of four differential equations, tracking the proportion of total individuals in each of four infected states (drug-sensitive untreated, drug-sensitive treated, drug-resistant untreated, drug-resistant treated), while the remaining individuals are uninfected (see **Methods**).

The basic reproductive ratio (R_0), defined as the average number of secondary infections caused by a single infected individual in an otherwise uninfected population, can be defined for each disease strain in this model and completely determines the equilibrium behavior (details in **Supporting Information**). If both infections have $R_0 < 1$, then neither strain of infection will persist and the equilibrium consists only of uninfected individuals. Otherwise, only the strain with the larger R_0 will persist (at a prevalence $1 - 1/R_0$, Figure 3B) while the other will go extinct. Based on the formulas for R_0, the resistant strain persists if and only if the coverage and efficacy of treatment are enough to offset the cost of resistance, $\varepsilon f > c$ (Figure 3C). Therefore, when the population is well-mixed, the heterogeneity introduced by partial treatment coverage ($0 < f < 1$) is insufficient to

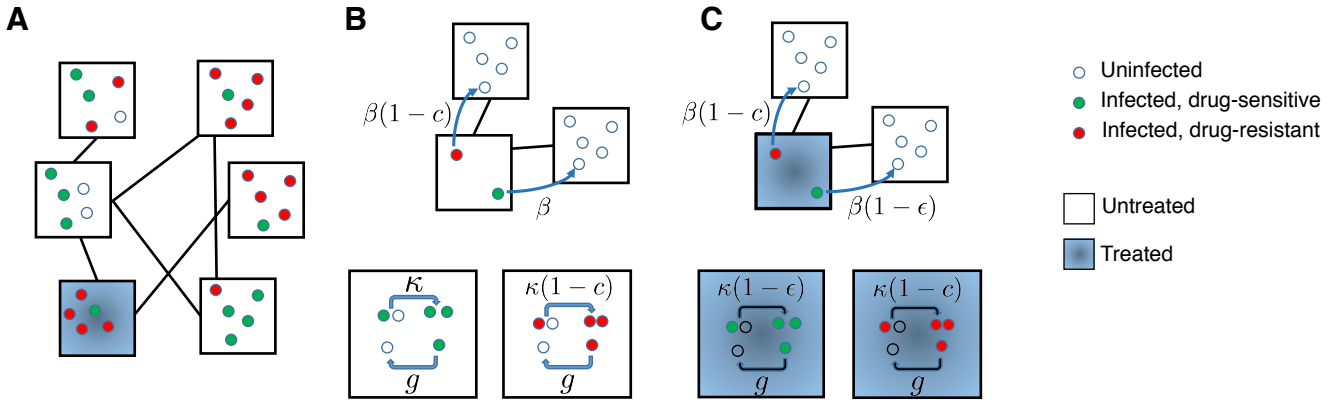

Figure 2. A structured population model for the spread of drug-resistant and drug-sensitive strains of an infection. **A)** An example population, divided into six equal subpopulations ("demes", black squares) of five individuals (circles). Infection can spread within demes, and also between demes that are connected (black lines). Individuals are categorized based on their infection status (uninfected: open circle, infected with drug-sensitive strain: green circle, infected with drug-resistant strain: red circle). The deme where an individual is located may determine whether or not they receive drug if infected (blue shading). **B)** Untreated deme: The wild-type strain (green) is transmitted at rate κ within a deme (bottom) and rate β between demes (top). Individuals infected with any strain recover with a rate g. The resistant strain (red) pays a cost c in its transmission rate with or without drug. **C)** Drug-treated deme: Transmission of the resistant strain is unaffected by treatment, but transmission of the wild-type strain (green) both within (bottom) and between (top) demes is reduced by a factor $(1-\varepsilon)$, where ε is the drug efficacy.

allow coexistence of wild-type and drug-resistant infections, and competitive exclusion always occurs, in agreement with other studies[24,28–30].

Limited coexistence in two connected subpopulations with heterogeneous treatment distribution

We next consider the case of two separate but connected subpopulations of equal size (Figure 3D). Upon infection, individuals in one of the demes will always receive treatment, while individuals in the other are never treated. The system is described by a set of equations tracking the proportion of individuals of each infection status (uninfected, infected with drug sensitive, infected with drug resistant) in each deme (treated vs untreated). The basic reproductive ratio can be derived for each strain (R_0^w and R_0^r) (see **Methods**).

This population supports qualitatively different infection dynamics than the single-deme case (Figure 3E-G). While most parameter regimes still lead to persistence of only the drug-sensitive or the drug-resistant strain (even when $R_0 > 1$ for both), there is now a stable equilibria where there is a mixture of both strains coexisting (Figure 3D). This region of coexistence occurs when the between-deme infection rate is relatively low compared to the within-deme rate ($\beta/\kappa < 0.7$ when $\varepsilon = 0.9$) and for intermediate values of the cost of resistance c. The mathematical derivations of these boundary regions and their physical intuition is included in the **Supplementary Information**. Interestingly, in this mixed equilibrium, both demes contain a mixture of sensitive and resistant strains, a stricter type of strain coexistence than a situation in which each strain dominates a different deme.

This simple example, consisting of only two preferentially-mixing subgroups of equal size, demonstrates that co-clustering of transmission and propensity for antibiotic use can lead to stable coexistence of sensitive and resistant strains at not only the population level but also the *sub*population level, suggesting a resolution for the ubiquity of coexistence in empirical data (point **(I)** in the Introduction). This host population structure produces a situation in which even though a strain is disfavored in a global sense by having a smaller R_0 value than its competitors , it can be locally favored and thus avoid extinction. However, in this simple example, the region in parameter space where coexistence is seen is relatively small, even for this extreme example of treatment clustering, motivating the question of whether more complex structures will expand the stability of coexistence. Moreover, this two-deme example cannot explain the observation that regions receiving the same amount of treatment or regions neighboring one another often support very different levels of resistance (observations **II** and **III**). Replicating real-world data requires finding situations in which neighbors with *the same* level of drug treatment still have very different levels of resistance. In the subsequent sections, we examine whether more complex population structures can recreate these observed patterns.

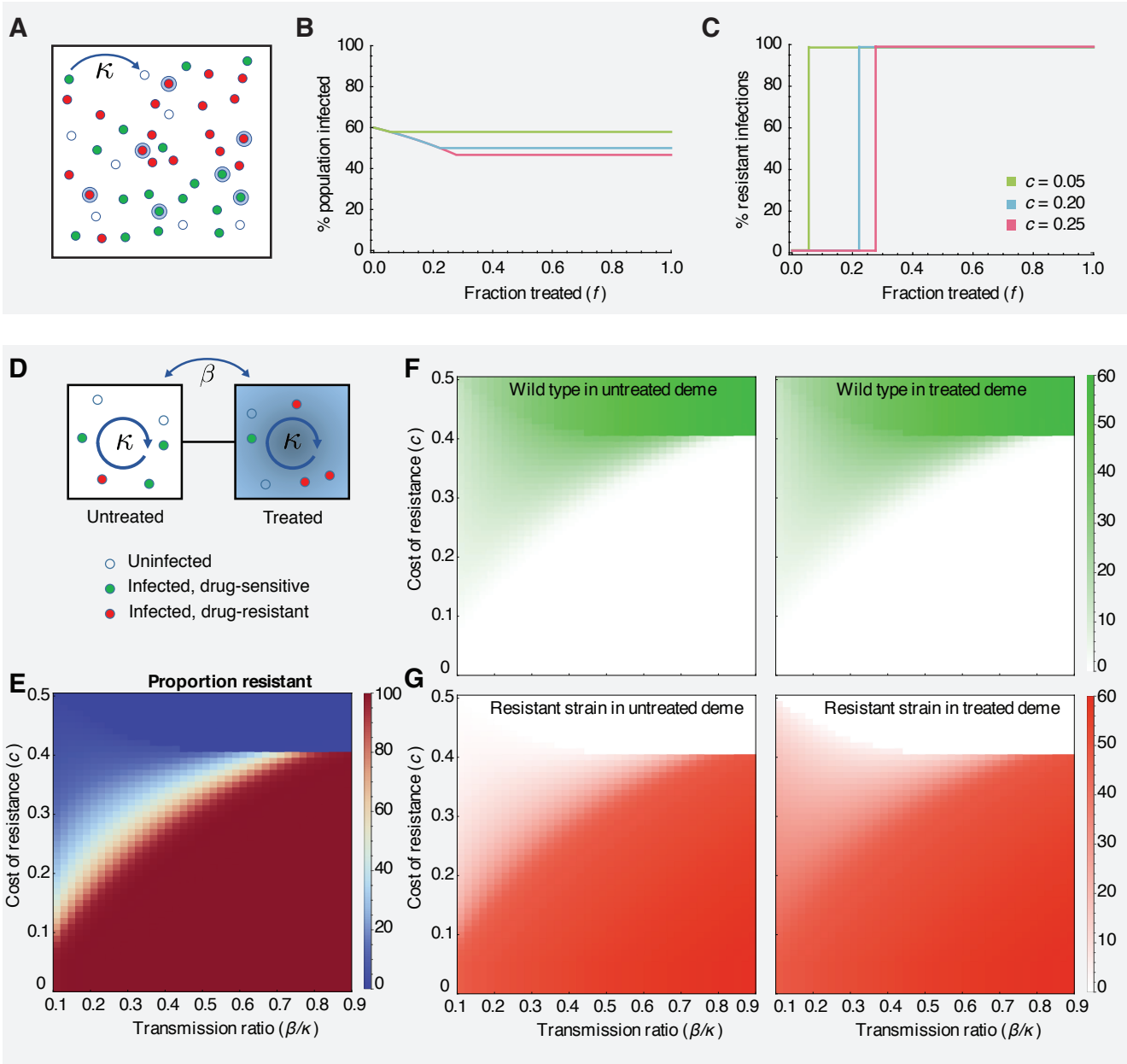

Figure 3. Competition between drug-sensitive and drug-resistant strains in a single well-mixed population (A-C) and two interconnected subpopulations (D-G). **A)** A population of individuals in a single well-mixed deme, in which a fraction f will receive drug treatment when infected (blue haloes). Individuals may be uninfected (hollow blue circles), or infected with either the wild-type (green circles) or drug-resistant (red circles) strain. **B)** The total prevalence of infection (wild-type + drug-resistant) as a function of the fraction of treated individuals (f) for different drug efficacies (ε) and costs of resistance (c). **C)** The % of infections that are drug-resistant as a function of the fraction of treated individuals (f) for different parameters. Infection switches between 0% and 100% resistant when $f = c/\varepsilon$. Coexistence never occurs. **D)** Schematic of a two-deme population (left-untreated, right-treated) and the two strains (green-wild type, red-resistant) considered in the model. **E-G)** Each panel shows the infection level (shading) as a function of the relative connectivity between demes (β/κ) and the cost of resistance (c). **E)** The % of all infections that are drug-resistant strain across the entire population. **F)** The % of individuals in each deme who are infected with the wild type strain. **G)** The % of individuals in each deme who are infected with the resistant strain. For all results, the transmission rate is $\kappa = 0.25$, the recovery rate is $g = 0.1$, and the treatment efficacy is $\varepsilon = 0.9$.

Robust coexistence in more complex population structures

Our results for simple one- and two-deme scenarios suggest that host population structure can promote coexistence of drug-sensitive and resistant strains of an infection. To investigate this relationship further, we extended our model to include an arbitrary number of demes and connectivity patterns (Figure 2A). For simplicity, we first assumed that each deme in the population was connected randomly to a fixed number of other "neighbor" demes, therefore creating a collection of random regular graphs[44] (Figure 4A). Although in reality much more complex human population structures are possible, these simplified networks provide a good base to examine the impact of specific properties on infection dynamics. Each deme was independently assigned to be "treated" (meaning if an individual in that deme became infected they would receive treatment, while individuals "untreated" demes would not).

Using these "meta-populations", we found that coexistence is possible for a broad range of costs of resistance, treatment levels, and population structures. As observed in data, the overall prevalence of resistance in a population is roughly correlated with the treatment coverage (Figure 4B). Moreover, coexistence also persisted at the level of individual demes. We defined 'robust' coexistence to mean that at least 80% of demes had at least 10% of infections caused by each strain. For the baseline parameter values we used, treatment levels between 35 and 55% supported robust coexistence in at least some population structures and treatment allocation schemes. For a lower cost of resistance ($c = 0.05$), robust coexistence occurred at a lower range of treatment levels (10% to 35%) (Figure 4C). When the amount of infection spread (β) arising from contact between "neighboring" demes increases relative to the within-deme spread (κ), the region of coexistence decreases (teal color). This is consistent with our findings that single well-mixed populations result in competitive exclusion of strains even for incomplete treatment.

These results show that even relatively simple host population structures can reconcile mathematical models with empirical observations of long-term coexistence between drug-sensitive and drug-resistant strains. These results also provide a connection between our model and the data for another spatio-temporal trend, **(II)**: the same overall rate of drug consumption can lead to different prevalence of drug resistance in different populations (even if treatment efficacy and cost of resistance are the same). For example, in Figure 4B, a particular treatment level (e.g. 20% for the red parameter set) can be associated with very different frequencies of resistance in different population structures (up 30% difference between graph 6 and 9, similar to the data in Fig. 1). While the number of neighbors of each deme does not seem to influence the likelihood of robust coexistence, it does effect the total population-wide resistance level (see Fig S2). A more detailed study of the role of higher-order network properties is presented in a later section.

Large differences in resistance levels are possible even between connected regions

With these complex multi-deme populations, we examined whether neighboring regions with equal levels of antibiotic consumption could sustain vastly different amounts of resistance (observation **(III)**. For each population structure, random treatment allocation, and parameter set, we chose all nearest-neighbor pairs which were both untreated, and generated a distribution of the pairwise differences in resistance levels between these neighbors (Figure 4D,E). We found that large differences in resistance levels between neighbors were common, with 34% of pairs receiving no treatment differing by more than 10% resistance prevalence, for our baseline parameter values when averaging across all population structures (Figure 4E). In the same simulations, 2% of pairs overall – and up to 15% in some structures – differed by greater than 30% resistance frequency, a value near the upper limit of observations from data shown in Figure 1. Differences of up to 60% between neighbors were observed in some individual simulations.

While populations with higher levels of between-deme mixing (higher β/κ, teal color) often supported persistence of both drug-sensitive and drug-resistant strains at equilibrium (Figure 4B), they did not support large difference in resistance levels between demes receiving the same treatment (Fig. 4E). In this parameter regime, sensitive and resistant strains are more segregated between untreated and treated demes, respectively, explaining why simulations on these structures rarely produced "robust" coexistence as we defined it (over 80% of demes supporting at least 10% of each strain) (Figure 4C). We found that populations with higher degrees of connection between demes rarely had pairwise differences in prevalence greater than 20% (Fig. S3). Overall, these findings corroborate the intuition that adding edges between demes or raising the between-deme infection rate of connected demes brings the system closer to well-mixed population and hinders coexistence.

Properties contributing to higher resistance

In order to understand how specific properties of host population structure contribute to the frequency of resistance within an individual deme and across a community of interconnected demes, we simulated infection dynamics on a large set of randomly generated structures. To generate a large ensemble of populations which varied in many graph-theoretic properties, we used the Watts-Strogatz[45] algorithm to create 1000 unique networks of 50 demes each, which had different average degree, variance in degree, centrality, efficiency, and mean path length (see **Methods** and Figure S4). For each graph, we generated 50 different allocations of treatment at overall coverage levels of 24% or 40%, which resulted in a wide range of resistance levels at equilibrium (Figure S5). Collecting the results for each graph-treatment allocation scenario, we employed LASSO (least

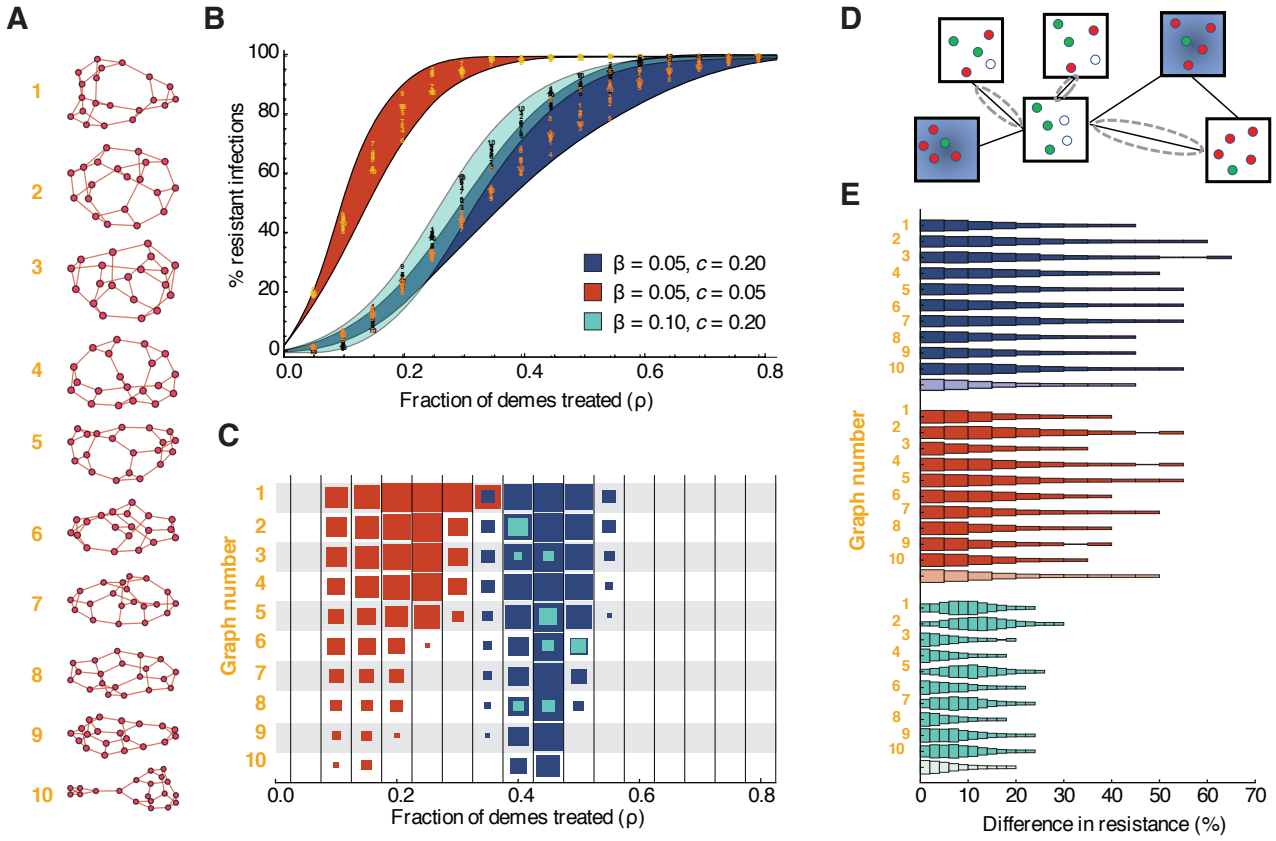

Figure 4. Dynamics of drug-resistant infections in populations consisting of networks of inter-connected demes. A) Randomly generated population structures on which infection was simulated. Each node represents a deme (a well-mixed sub-population of individuals), and each edge indicates that infection can spread in either direction between those two demes. Ten example populations, each with twenty demes randomly connected to three neighbors each, were indexed out of 1000 total simulated to represent a broad range of outcomes. **B)** Fraction of infections that are resistant in the entire population (y-axis) versus fraction of demes treated, ρ (x-axis). Each color represents a different parameter set (blue background — baseline, red background — lower cost of resistance, teal background — more between-deme connectivity). Numbers show data points for the ten example populations. The colored envelope is created by shading between sigmoidal curves that encompass all the data. **C)** For each population structure shown (y-axis) and each treatment level (x-axis), the proportion of simulations that resulted in *robust* coexistence between drug-sensitive and drug-resistant strains is shown (by the colored area of the box). Robust coexistence was defined as at least 80% of demes supporting both strains at frequencies above 10%. **D)** Differences in resistance levels (% of all infections that are with the drug resistant-strain) are measured between all pairs of directly-connected untreated demes. **E)** Histograms showing the distribution of pairwise differences in resistance for a given population structure. Lighter shaded histograms combine results from all population graphs. All simulations used kinetic parameters $\kappa = 0.25$, $g = 0.1$, and $\varepsilon = 0.9$, and pooled results from 100 simulations with different random allocation of treatment across demes. Pairwise differences were calculated with 30% treatment.

Regression on deme properties			Regression on graph properties		
Rank	Property	Associated with	Rank	Property	Associated with
1	Treated	↑ **resistance** (+/+)	1	U-U-T	↓ **coexistence** (-/+)
2	Degree	↓ **coexistence** (-/+)	2	U-T	↑ **resistance** (+/+)
3	% treated 1 edge away	↑ **resistance** (+/+)	3	T-T	↑ **resistance** (+/+)
4	% treated 2 edges away	↑ **resistance** (+/+)	4	U-T-T	↑ **resistance** (+/+)
5	Eigencentrality	↓ **coexistence** (-/+)	5	T-T-T	↑ **coexistence** (+/-)
6	% treated 3 edges away	↑ **resistance** (+/+)	6	Local efficiency	↓ **coexistence** (-/+)
			7	Clustering coefficient	↓ **resistance** (-/-)
			8	Global efficiency	↓ **coexistence** (-/+)
			9	Mean path length	↑ **coexistence** (+/-)
			11,12	Mean,variance of degree	—, —

Table 1. Properties of demes and graphs that accurately and sparsely predict the frequency of drug-resistant infections according to LASSO regression. The shorthand with "U" and "T" indicates the frequency of pairs and triplets of **u**ntreated and **t**reated demes; thus, U-T-T indicates the frequency of three connected demes of which two are drug-treated, normalized by the number of all three-deme combinations in the graph. LASSO regression was conducted separately for two levels of treatment (24% demes treated, leads to $< 50\%$ resistant, and 40% demes treated, leads to $> 50\%$ resistant). The rank of each property is determined from the magnitude of the predictor coefficient at the value of the regularization constraint minimizing the mean-squared error after five-fold cross-validation (Fig. S6, S7), averaged over both the low and high treatment case. For each regression, the sign of the direction of the relationship between the property and the frequency of resistance was determined (+ or -), and used to summarize the associations: "↑ resistance": (+/+) the property is always associated with increased resistance , "↓ resistance": (-/-) the property is always associated with decreased resistance , "↑ coexistence": (+/-) the property is associated with increased resistance when resistance is rare (low treatment) but decreased resistance when resistance is common (high treatment) , "↓ coexistence": (-/+) the property is associated with decreased resistance when resistance is rare (low treatment) but increased resistance when resistance is common (high treatment), "—": the property is not used in the "best" (error-minimizing) model.

absolute shrinkage and selection operator) regression to identify the most important factors contributing to the frequency of resistant infections (Table 1).

We first examined predictors of the resistance level within a deme. The most important contributor to resistance was whether or not the deme itself was treated, and the distribution of treatment among the deme's first, second and third nearest neighbors were also significant – albeit less important – predictors. The next strongest predictors was the total number of first degree neighbors of a deme (the "degree"). Higher degree is associated with lower resistance when resistance is rare (at 24% treatment), but higher resistance when it is common (at 40% treatment), and hence acts against coexistence ($-/+$ in Table 1). The same is true for the eigencentrality of the deme, a measure of its overall level of connectivity within the network. These results imply that highly-interconnected, hub-like demes are more likely to harbor the more common strain, while demes that are on the periphery of a network are more likely to support the strain that is rarer in the population overall. Hence, when resistance is rare, it may be preferentially found in less connected demes, whereas when it is more common, it may cluster in more central demes.

We next evaluated predictors of resistance at the level of the entire population. Surprisingly, the simplest statistics of a graph – including the mean and variance in the number of neighbors of each deme – were the least important contributors to resistance. Rather, the best predictors involved a mix of treatment-clustering properties, like the proportion of sets of three connected demes where two of the demes were treated ("U-T-T"), and node-clustering properties, such as measures of "efficiency". Treatment promotes higher resistance levels when it is distributed more evenly throughout the population (e.g. "U-T-T", "U-T"), but only the isolation of communities receiving the most treatment ("T-T-T") facilitates coexistence. This suggests that the effect of increasing antibiotic consumption on resistance levels will depend strongly on which communities experience the increases. The clustering coefficient and local efficiency are both measures of how interconnected the neighbors of a given deme are, on average. Global efficiency and mean path length are both measures of the ease of moving between any two random demes in the whole population. Coexistence is promoted by higher mean path length, which leads to graphs with more segregated transmission clusters, but hindered by global and local efficiency. The more interconnected a network is, the harder it is for multiple strains to coexist. The reason why the clustering coefficient always acts against resistance is not entirely clear, but its interpretation is complicated by the fact that the treatment-clustering properties are partially determined by the level of

clustering in the graph.

Robustness of results to mechanism of treatment action

Throughout this paper, we have modeled the effect of antibiotic treatment as reducing the ability of an infected individual to spread infection to others. Hence, treated individuals are still infected, and therefore immune to infection with the other strain, but do not contribute to transmission (or contribute less, if treatment is imperfect, $\varepsilon < 1$). Alternatively, treatment could instead accelerate the rate at which an individual recovers from infection, a mechanism that has been used in many previous models of antibiotic resistance[28,29]. To determine if the mechanism of treatment action had any influence on our findings about the spatiotemporal patterns of resistance and to facilitate comparison of our results with those of other models, we considered a variant of the model where in treated individuals, the recovery rate increases from $g \to g + \tau$ (see **Supplementary Methods** for details). We found that all of the results reported in the main text are recapitulated when treatment increases recovery rate, including the absence of coexistence in partially-treated well-mixed populations (Fig. S8), a limited though larger parameter regime of coexistence in two-deme populations (Fig. S9), common coexistence in multi-deme populations (Fig. S9) and divergent resistance levels between neighboring demes receiving the same treatment level (Fig. S11).

Discussion

The prevalence of antibiotic-resistant bacterial infections displays several surprising spatial and temporal trends that defy canonical predictions of infection dynamics. The goal of this paper was to create a simple and flexible model for competition between drug-sensitive and drug-resistant strains of an infection in a structured host population and examine whether it could capture several of these trends. The first perplexing pattern observed in the data is the long-term coexistence of both sensitive and resistant strains at intermediate levels, which many other studies have attempted to explain[24,28–30]. Our simulations reproduce coexistence for a wide range of parameter values and in fact suggests that it is a natural feature to be expected in spatially (or otherwise) heterogeneous populations with competing infection. In addition, we also captured more nuanced trends observed in surveillance data that other models have ignored. Regions that administer similar amounts of a particular antibiotic may experience persistent differences in drug resistance levels for many years, and these "frozen gradients" are observed even if regions are neighboring one another. These phenomena were robustly reproduced even in our simple model (Figure 5A). We found that in a collection of partially-mixing subpopulations, subtle details about the pattern of connectivity and how treatment is distributed between them can lead to large differences in resistance prevalence between regions that otherwise seem similar. We were able to identify specific covariates that contribute to resistance levels and to promoting coexistence with drug-sensitive strains. Similar to the use of "risk-mapping" to predict the distribution of vector-borne diseases based on environmental[46–49] and human[49–53] factors, these findings suggest that spatial risk-based assessment may be useful for the study of drug resistance. Surveillance on a finer scale will be needed to test these ideas on empirical data.

Increasing levels of antibiotic resistance are widely considered to be a major public health threat, and attempts are continually underway to forecast resistance levels with or without additional interventions using at least some type of mathematical modeling. Our findings highlight the fact that infection models which rely on assumptions of well-mixed populations are unlikely to be a useful tool for this task. As these models cannot reproduce ubiquitous trends observed in existing data on drug resistance epidemiology, they are unlikely to make predictions that are even qualitatively trustworthy (Figure 5B). For example, it is very important to know whether resistance for specific bug-drug combinations will tend towards 100% or settle at an intermediate level, and so any model must be able to explain the phenomena of coexistence for most strains. While clearly any degree of resistance complicates clinical management, the economic impact of partial vs complete resistance in a population could differ dramatically. In addition, by ignoring the details of spatial or demographic heterogeneity, we may miss an opportunity to more effectively target intervention or surveillance to particular subpopulations. The obvious difficulty is that such models will may have to be highly-tailored towards a particular disease, and currently, there is a stark lack of data required to construct or calibrate realistic spatial models. More routine high-resolution mapping of resistance levels over space and time[35], as well as data to estimate population connectivity patterns over which disease spreads and patterns of antibiotic consumption[54,55], are sorely needed.

Throughout this paper we have considered a relatively simple type of population structure, while in reality it could be much more complicated. There may be multiple levels of structure - from a household to countries - with much more complicated patterns of connectivity and a continuum of transmission rates. Structures may be dynamic, as movement and interaction patterns of humans, animals, or disease-carrying material change over time. The mechanism leading to population structure varies depending on the infection of interest. For hospital-acquired infections, the movement of medical staff or the spatial arrangement of patients may be most important, while for community-acquired infections, the nature of close household, school, or workplace contacts may be of primary interest. For sexually-transmitted infections, networks of sexual contacts determine disease spread, while for diseases of livestock, relatively well-mixed contact within crowded farms along with patterns of transfer between farms may create an overall "meta-population" structure. Host population structures are notoriously difficult to

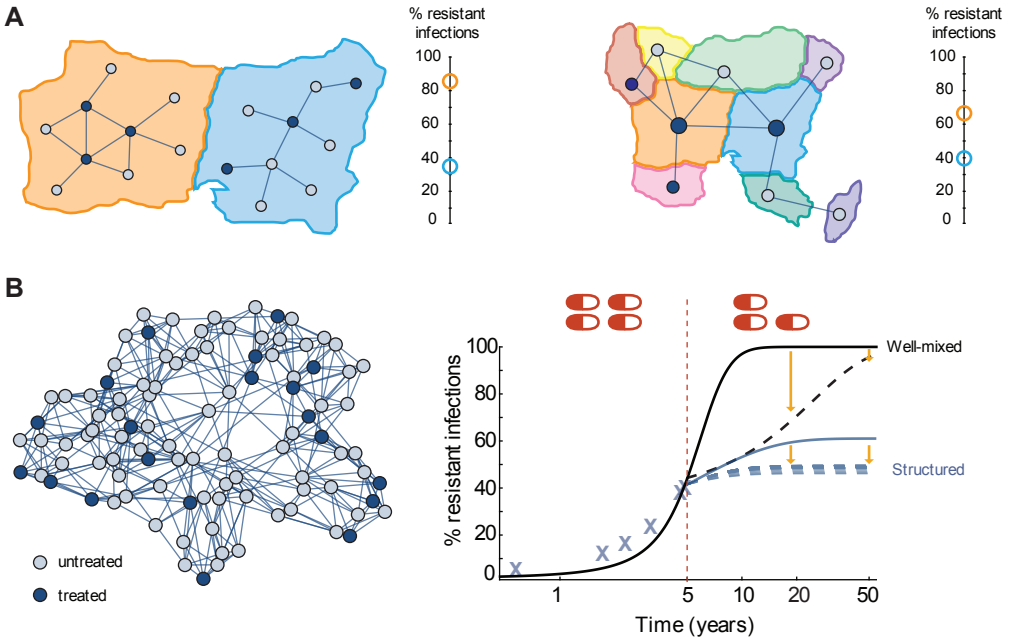

Figure 5. Implications for interpreting and predicting resistance levels. A) Our analysis suggests two possible explanations for large differences in resistance levels between two regions (orange and blue) receiving similar amounts of antibiotics. **Left:** The regions may differ in the distribution of antibiotic consumption within the population and its connection with the underlying transmission network. **Right:** The regions may differ in their connectivity to other regions which consume different amounts of antibiotics. **B)** Predictions about the impact of interventions on future resistance levels can be incorrect if population structure isn't accounted for. **Left:** A population consisting of 100 sub-populations, with 20% treatment coverage at baseline. **Right:** Simulations of resistance levels in the "true" structured population (blue) with and without a hypothetical intervention (reduction to 15% treatment) applied at year 5. For comparison, predictions of a well-mixed model that with the same parameters (black), which approximates the dynamics before year 5 but diverges afterwards. The well-mixed model would predict large reductions in resistance in the short term but eventual fixation of resistance nonetheless (black dashed line), whereas the structured model predicts modest but sustained reductions which depend on the particular sub-populations targeted (blue dotted lines), and long-term coexistence. All simulations used kinetic parameters $\kappa = 0.25$, $\beta = 0.05$ $g = 0.1$, and $c = 0.2$. For the left of (A) and for (B), $\varepsilon = 0.9$, and for the right of (A), $\varepsilon = 0.5$. The population in (B) is a Watts-Strogatz network with degree 4 and re-wiring probability 0.1.

measure in practice, but some approximation of them is likely needed to create realistic models for the spread of drug resistant infections. Technological advancements that allow better tracking of human mobility, such as wearable proximity censors[56–58] or fitness trackers, cell phone location data [59–62], or large-scale transit networks[63,64], provide one source of population structure data. Advances in genetic epidemiology allow us to retrospectively trace the spread of infection between individual hosts or communities using pathogen phylogenies[65–69], which also provides information on population structure.

Spatial heterogeneity has been understood by population geneticists to be an important factor in the spread of genes since the days of Fisher[70] and later Kimura[71]. Other classic work uncovered mechanisms by which the coexistence — as opposed to competitive exclusion — of species could be promoted by spatial heterogeneity[72,73]. More recently, Gavrilets & Gibson[74] examined a two-deme system with two competing types, each of which has higher fitness than the other in one of the two demes, and observed a phase diagram very similar to that for our two-deme infection model (Fig. 3), in which polymorphic equilibria were possible. This idea was later extended to multiple connected demes[75], where it was demonstrated that spatial heterogeneity can preserve a species in spite of fitness differences which in a well-mixed model would drive it to extinction. Other work involving more complex population structures[76] or more complex fitness distributions across space[77,78] has examined how the rate of invasion of rare mutants depends on spatial heterogeneity. This previous work all involved traditional population-genetic models, such as Moran or Wright-Fisher processes, and before this study it was not clear which results would generalize to infection dynamics models. For example, the role of network structure in modulating the fixation probability of new strains differs dramatically between Moran processes[76] and infection models[79].

While our results show that population structure can explain observed spatiotemporal patterns of antibiotic resistance, it is unlikely be the only mechanism contributing to either long-term coexistence or differences between regions. Persistence of both drug-sensitive and drug-resistant strains is also promoted by confection and superinfection[28–30], by overlaps between resistance status and serotype[28,29] or other traits under balancing selection[33], and by heterogeneities in transmission or recovery rates as opposed to treatment coverage[28,29]. Regions receiving similar overall levels of drug but experiencing different frequencies of resistant infections could alternatively be explained by local differences in i) prescribed vs consumed doses, ii) who receives drugs (e.g. age group, hospital vs outpatient) and for what type of condition, iii) genetic background of the pathogen and the cost of resistance, or iv) transmission or recovery rates. It is difficult to judge precisely how realistic our results are, since our models of infection dynamics and population structure are both dramatically simplified, but it likely cannot explain coexistence over the entire range of parameters for which it is observed empirically. In addition, in our results it is relatively rare to see large differences in resistance prevalence between regions with similar treatment rates unless they have very different connectivity patterns within or between them, and unless there are large spatial or demographic heterogeneities in treatment use that overlap with clustering of disease transmission. In reality, multiple mechanisms probably act in concert to explain the observed patterns. Moreover, there is unlikely to be single explanation for each different bug-drug pair for which these trends are observed. Other disease-specific factors that may be relevant to the ecology of antibiotic resistance include whether the bacteria is primarily pathogenic or also a commensal colonizer, whether infection is more common in the community or in healthcare facilities, whether resistance is carried on plasmids or chromosomally, and whether there are environmental reservoirs and if they are exposed to antibiotics.

Methods

Model assumptions

Our two-strain model of infectious disease dynamics makes a number of simplifying assumptions. We assume that individuals can transmit the infection immediately upon becoming infected, hence ignoring any sort of "latent" phase or "exposed" class. Birth, death, and migration into and out of the population are not modeled. The infection is assumed to be non-lethal. Infected individuals all recover to be susceptible again, which presumes that there is no long-lasting immunity to re-infection. We do not allow for any co-infection (simultaneous infection with both strains) or super-infection (replacement of one strain within an individual by another), which implies that there is complete cross-immunity during infection. We model resistance as a binary trait: there are only two strains, either completely sensitive to the drug, or completely resistant. In reality, there may be strains with intermediate levels of resistance. Individuals who are assigned a status of "treated" will receive treatment immediately upon being infected, every time they are infected. This ignores the possibility that there may be a delay to receiving treatment or that treatment behavior may vary within an individual over time. Treatment is assumed to prevent transmission of the infection, but may not be perfectly effective at doing this (which could approximate the effect of treatment delay or imperfect drug efficacy). In the **Supplementary Methods**, we consider an alternative model in which treatment acts to increase the rate of recovery from infection instead.

Our population structure consists of M subpopulations ("demes") of equal size D for a total population of size N. The structure is assumed to be static. For each model, the transmission rates κ and β were scaled so that the model behavior did not depend on N, D, or M. In the one- and two-deme case, κ was multiplied by N, and in the multi-deme cases, both κ and β were multiplied by D. A more detailed derivation of this scaling is given in the **Supplementary Methods**. This scaling is equivalent

to saying that individuals are limited in how many contacts they can have with others, and that this limit is independent of the total population size.

By using differential equations to model infection, we assume that the population sizes of individual demes are large enough so that variation from the average behavior is not important. We additionally assume that resistance is pre-existing and that stochastic extinction of either strain from the population is not possible. For the most part, we examine only the equilibrium behavior of the models.

Model equations

Well-mixed population (single deme): Infection spreads between all individuals with the within-deme transmission rate κ. A fraction f of all individuals are treated with drug immediately upon infection by either strain. We track the proportion of the total population who are infected with the wild type and untreated (w_u), infected with the wild type and treated (w_t), infected with the resistant strain and untreated (r_u), and infected with the resistant strain and treated (r_t). The number of uninfected (susceptible) individuals is given by $s = 1 - w_u - w_t - r_u - r_t$. The dynamics are then described by the following set of differential equations:

$$\begin{aligned}
\dot{w_u} &= \kappa(w_u + (1-\varepsilon)w_t)((1-f) - w_u - r_u) - gw_u, \\
\dot{r_u} &= \kappa(1-c)(r_u + r_t)((1-f) - w_u - r_u) - gr_u, \\
\dot{w_t} &= \kappa(w_u + (1-\varepsilon)w_t)(f - w_t - r_t) - gw_t, \\
\dot{r_t} &= \kappa(1-c)(r_u + r_t)(f - w_t - r_t) - gr_t.
\end{aligned} \tag{1}$$

Two deme population: Infection spreads between all individuals in the same deme with rate κ, and between any two individuals in different demes at rate β. The demes are of equal size D. Treatment is assigned to all individuals within only one of the two demes. We track the proportion of individuals in the treated deme that are infected with the wild-type/drug-sensitive strain (w_u) and infected with the resistant strain (r_u), and the same proportions for the treated deme (w_t and r_t). The number of uninfected (susceptible) individuals in the untreated deme is $s_u = 1 - w_u - r_u$ and in the treated deme is $s_t = 1 - w_t - r_t$. The system is then described by the following set of differential equations (see **Supporting Information** for a derivation and generalization to unequal deme sizes):

$$\begin{aligned}
\dot{w_u} &= (1/2)(\kappa w_u + \beta(1-\varepsilon)w_t)(1 - w_u - r_u) - gw_u, \\
\dot{r_u} &= (1/2)(1-c)(\kappa r_u + \beta r_t)(1 - w_u - r_u) - gr_u, \\
\dot{w_t} &= (1/2)(\beta w_u + \kappa(1-\varepsilon)w_t)(1 - w_t - r_t) - gw_t, \\
\dot{r_t} &= (1/2)(1-c)(\beta r_u + \kappa r_t)(1 - w_t - r_t) - gr_t.
\end{aligned} \tag{2}$$

Multi-deme population: Infection spreads between all individuals in the same deme with rate κ, and between any two individuals in different but connected demes at rate β. There are M demes each of size (D). Each deme may only be able to spread infection to a sub-set of other demes, and the connectivity of the population is described by the adjacency matrix Δ_{ij} ($\Delta_{ij} = 1$ if an individual in deme j can be infected by an individual in deme i). The *degree* of deme i, d_i, is the number of neighbors it is connected to. Treatment is assigned at the level of the deme, described by indicator variable T_i (where $T_i = 1$ if deme i is drug-treated and $T_i = 0$ otherwise). The fraction of demes that are treated is ρ. The system of equations describing the fraction of individuals in each deme who are infected, with either the wild-type (w_i) or drug-resistant strain (r_i) is

$$\begin{aligned}
\dot{w_i} &= \kappa(1-\varepsilon T_i)w_i(1 - w_i - r_i) + \beta(1 - w_i - r_i)\sum_j \Delta_{ji}(1-\varepsilon T_j)w_j - gw_i, \\
\dot{r_i} &= \kappa(1-c)r_i(1 - w_i - r_i) + \beta(1-c)(1 - w_i - r_i)\sum_j \Delta_{ji} r_j - gr_i.
\end{aligned} \tag{3}$$

Derivations for the equilibria, stability conditions, and basic reproductive ratio (using the next-generation technique[80]) are given in the **Supplementary Information**.

Numerical results

The differential equations were numerically integrated in MATLAB, using the Runge-Kutta solver provided in the function `ode45`. The initial condition for each parameter-population structure combination was each deme having a level of infection with the wild-type that would occur if there were no resistant strain (a fraction $1 - 1/R_0$ of individuals are infected with the wild type), and a very small level of infection with the resistant strain (10^{-3}). The rest of the population was uninfected.

The system was integrated until an equilibrium was reached, which was defined as the point at which the time derivatives for all strains in all demes changed by less than 10^{-4} per day. To ensure that the equilibrium reached by the solver is stable,

once an equilibrium was reached by the above standard, we applied a small random fluctuation with magnitude not exceeding 1% to the value of all strains in all demes. If the same equilibrium was arrived at after repeating this process three times, the equilibrium was taken to be sufficiently stable and used as the result of that simulation. To determine whether there was only a single positive and stable equilibrium for the multi-deme system for a given structure and parameter set, we simulated the same systems many times with uniformly-at-random initial conditions. For each deme, the initial fractions $[w_i, r_i]$ were both drawn from the uniform distribution on $[0, 0.7]$. If their sum exceeded 1, both were re-drawn. We found that the same equilibrium was always reached (Figure S12). More details are given in the **Supplementary Information**.

Network generation

Random regular graphs: Random regular graphs are randomly-generated graphs in which every node has the same degree (number of incident edges). There are multiple existing algorithms to create random regular graphs. Our random regular graphs were generated according to a pairwise-construction mechanism[81] implemented in MATLAB[82].

Watts-Strogatz graphs: The Watts-Strogatz algorithm is a method of constructing networks with M total nodes and $M * d$ total edges, but with heterogeneous properties. In the original construction, a ring lattice is formed with each of the M nodes connected with d edges to neighbors in a ring formation. Then, with probability p, the target of each edge is rewired to a uniformly random node. For our 1000 graphs, we selected $M = 50$, but allowed degree d to be selected uniformly-at-random for each graph from 3, 4, 5, or 6, and allowed p to be selected uniformly-at-random for each graph from $[0.1, 0.6]$. The resulting graphs have a broad distribution all graph properties of interest (Figure S4) and of the frequency of resistance (Figure S5).

Network statistics

A set of graph-theoretical properties were tested for their relationship to resistance levels within demes and within populations. A description of these properties is given below.

Eigencentrality: For a node i, its eigencentrality is a measure of not only its centrality (having a high degree, for example), but is also weighted according to how high the centrality scores of its neighbors are; as an example, an academic paper with many citations may have high centrality, but if the papers citing it have themselves low centrality then the eigencentrality of the original paper may be low. This is written as an eigenvector-type equation for the eigencentrality E_i: $E_i = (1/\lambda)\sum_j \Delta_{ij} E_j$, where λ is a normalization factor.

Global efficiency: In general, the average efficiency of a graph is defined as $\mathrm{Eff}(G) = \frac{1}{M(M-1)}\sum_{i \neq j} d_{ij}^{-1}$, where M is the number of demes/nodes and d_{ij} is the smallest number of edges between nodes i and j. Therefore, graphs that have higher efficiency have on average shorter path lengths between nodes. Since $M = 50$ is fixed in all of our structures examined in this section, the only varying factor is the distances between nodes; the smaller the distances between all given nodes, the higher the efficiency. Global efficiency involves comparing the graph to the most efficient possible graph with as many nodes. Since this simply involves another prefactor which is equal for all of our graphs, our global efficiency is equal to the average efficiency.

Local efficiency: Local efficiency is defined as $\mathrm{Eff}_l(G) = M^{-1}\sum_i \mathrm{Eff}(G_i)$, where G_i is the subgraph in which we keep the neighbors of node i but remove i itself, and the efficiency is defined as above.

Clustering coefficient: The network average clustering coefficient was calculated by averaging the local clustering coefficient over all nodes. The local clustering coefficient is calculated for a node i as the fraction of all possible edges that could exist between the neighbors of i which actually do exist in the network. Said another way, the local clustering coefficient is determined by counting all triplets (sets of three connected nodes) centered on node i and then calculating the fraction of these triplets that are triangles (all three nodes connected to each other).

Mean path length: Given any two nodes, we can calculate the path length as the smallest number of edges that must be traversed to move from one to the other. The average path length is the average over the distribution that arises from calculating this smallest number for all possible pairs of nodes.

LASSO regression

LASSO[83] is an advanced regression technique which assigns not only a regression coefficient, but also a hierarchical importance to each covariate. This is accomplished by forcing the sum of all the regression coefficients to be less than a certain number; in optimal fitting, some coefficients are then set to zero (i.e., deemed unimportant) rather than assigned a non-zero value. This is usually accomplished by adding the sum of all the regression coefficients, multiplied by a regularization parameter λ, to the quantity being minimized (such as sum of the squared error). When λ is large, having few covariates in a model becomes more important than the model fitting the data accurately. Therefore, the error-minimizing model usually occurs for some intermediate value of λ.

We employed Matlab's `lasso` (and `lassoPlot`) function (which imposes the LASSO constraint on the $\mathbf{L}_1$ norm of all fit coefficients onto a linear regression) with the various graph (and node) properties as possible model variables. The target to fit was either the rate of resistant infections (when considering an entire population of demes) or simply the prevalence of resistant infections (within a single deme).

Although a rigorous and robust significance test for the LASSO is still lacking, we can compute the mean-squared error of the predictions of the model to the actual data for cross-validation. Here we choose five-fold validation, meaning that the data is randomly split into five chunks of equal size. Four of these chunks are used to fit the regression, and then the mean-squared error between the actual fifth chunk and its prediction by the regression is measured. If this mean-squared error is sufficiently small, the model is a good fit. The results of this cross-validation procedure for LASSO regression fits to both deme and population-level properties are shown in Fig. S6.

Acknowledgments

We thank Jeff Gerold, Sam Sinai, Kamran Kaveh, Anjalika Nande, Gabriel Leventhal, and participants in the 2018 Gordon Research Conference on Drug Resistance for helpful discussions. **Funding:** This work was supported by the National Institutes of Health (DP5OD019851) and the Bill & Melinda Gates Foundation (OPPOPP1148627). **Author Contributions:** M.S.K and A.L.H. conceived the study. M.S.K., M. A. N, and A.L.H. analyzed the model. M.S.K. performed all simulations, network analysis, regression, and stability analysis of the model and analyzed the resistance data. M.S.K., M. A. N, and A.L.H. interpreted the results and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data availability:** All data shown is available from the original source at the links provided in the references. All equations and algorithms needed to generate the results of this paper are described in the Methods/Supplementary Methods, and code to simulate the model equations is available from the authors by request.

References

1. Roberts, R. R. *et al.* Hospital and societal costs of antimicrobial-resistant infections in a chicago teaching hospital: Implications for antibiotic stewardship. *Clin. Infect. Dis.* **49**, 1175–1184 (2009).
2. McDonald, L. C. Trends in antimicrobial resistance in health care: Associated pathogens and effect on treatment. *Clin. Infect. Dis.* **42**, S65–S71 (2006).
3. Unemo, M. & Nicholas, R. A. Emergence of multidrug-resistant, extensively drug-resistant and untreatable gonorrhea. *Futur. Microbiol.* **7**, 1401–1422 (2012).
4. Neuner, E. A. *et al.* Treatment and outcomes in carbapenem-resistant klebsiella pneumoniae bloodstream infections. *Diagn. Microbiol. Infect. Dis.* **69**, 357 – 362 (2011). DOI https://doi.org/10.1016/j.diagmicrobio.2010.10.013.
5. Perez, F. *et al.* Global challenge of multidrug-resistant acinetobacter baumannii. *Antimicrob. Agents Chemother.* **51**, 3471–3484 (2007). DOI 10.1128/AAC.01464-06.
6. Reardon, S. WHO warns against 'post-antibiotic' era. *Nat. News* (2014). URL http://www.nature.com/news/who-warns-against-post-antibiotic-era-1.15135. DOI 10.1038/nature.2014.15135.
7. Molteni, M. The Post-Antibiotic Era Is Here. Now What? *Wired* (2017). URL https://www.wired.com/story/the-post-antibiotic-era-is-here-now-what/.
8. Anderson, R. & May, R. *Infectious Diseases of Humans — Dynamics and Control* (OUP, Oxford, UK, 1992).
9. Opatowski, L., Guillemot, D., Boëlle, P.-Y. & Temime, L. Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. *Curr. Opin. Infect. Dis.* **24**, 279 (2011). URL https://journals.lww.com/co-infectiousdiseases/fulltext/2011/06000/Contribution_of_mathematical_modeling_to_the_fight.18.aspx. DOI 10.1097/QCO.0b013e3283462362.
10. O'Neill, J. The Review on Antimicrobial Resistance: Final Report and Recomendations. Tech. Rep. (2016). URL https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf.
11. Kraker, M. E. A. d., Stewardson, A. J. & Harbarth, S. Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? *PLOS Medicine* **13**, e1002184 (2016). URL http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002184. DOI 10.1371/journal.pmed.1002184.
12. Liñares, J., Ardanuy, C., Pallares, R. & Fenoll, A. Changes in antimicrobial resistance, serotypes and genotypes in *Streptococcus pneumoniae* over a 30-year period. *Clin. Microbiol. Infect.* **16**, 402–410 (2010).
13. European Centre for Disease Prevention and Control. Antimicrobial resistance surveillance in Europe 2015. *Annu. Rep. Antimicrob. Resist. Surveillance Netw. (EARS-Net).* **Stockholm: ECDC** (2017).
14. Goossens, H., Ferech, M., Vander Stichele, R., Elseviers, M. & ESAC Project Group. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet (London, England)* **365**, 579–587 (2005). DOI 10.1016/S0140-6736(05)17907-0.

15. Swiss Centre for Antibiotic resistance. Interactive database - antibiotic resistance. URL http://www.anresis.ch/index.php/Interactive-database.html.

16. Fenoll, A., Bourgon, C. M., Muñóz, R., Vicioso, D. & Casal, J. Serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* isolates causing systemic infections in spain, 1979-1989. *Rev. Infect. Dis.* **13**, 56–60 (1991).

17. Fenoll, A., Jado, I., Vicioso, D., Pérez, A. & Casal, J. Evolution of *Streptococcus pneumoniae* serotypes and antibiotic resistance in spain: Update (1990 to 1996). *J. Clin. Microbiol.* **36**, 3447–3454 (1998).

18. Bergman, M., Nyberg, S. T., Huovinen, P., Paakkari, P. & Hakanen, A. J. Association between antimicrobial consumption and resistance in escherichia coli. *Antimicrob. Agents Chemother.* **53**, 912–917 (2009).

19. Song, J.-H. *et al.* High prevalence of antimicrobial resistance among clinical streptococcus pneumoniae isolates in asia (an ansorp study). *Antimicrob. Agents Chemother.* **48**, 2101–2107 (2004).

20. Camargos, P., Fischer, G. B., Mocelin, H., Dias, C. & Ruvinsky, R. Penicillin resistance and serotyping of streptococcus pneumoniae in latin america. *Paediatr. Respir. Rev.* **7**, 209 – 214 (2006).

21. Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? *Curr. Opin. Microbiol.* **9**, 461–465 (2006).

22. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat. Rev. Microbiol.* **8**, 260–271 (2010).

23. Melnyk, A. H., Wong, A. & Kassen, R. The fitness costs of antibiotic resistance mutations. *Evol. Appl.* **8**, 273–283 (2014).

24. Spicknall, I. H., Foxman, B., Marrs, C. F. & Eisenberg, J. N. S. A modeling framework for the evolution and spread of antibiotic resistance: Literature review and model categorization. *Am. J. Epidemiol.* **178**, 508–520 (2013).

25. Gauze, G. Experimental studies on the struggle for existence. *J. Exp. Biol.* **9**, 389–402 (1932).

26. Hardin, G. The competitive exclusion principle. *Sci.* **131**, 1292–1297 (1960).

27. Hutchinson, G. E. The paradox of the plankton. *Amer. Nat.* **95**, 137–145 (1961).

28. Colijn, C. *et al.* What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of Streptococcus pneumoniae? *J. The Royal Soc. Interface* rsif20090400 (2009). URL http://rsif.royalsocietypublishing.org/content/early/2009/11/20/rsif.2009.0400. DOI 10.1098/rsif.2009.0400.

29. Cobey, S. *et al.* Host population structure and treatment frequency maintain balancing selection on drug resistance. *J. The Royal Soc. Interface* **14**, 20170295 (2017). URL http://rsif.royalsocietypublishing.org/content/14/133/20170295. DOI 10.1098/rsif.2017.0295.

30. Davies, N. G., Flasche, S., Jit, M. & Atkins, K. E. Within-host dynamics explain patterns of antibiotic resistance in commensal bacteria. *bioRxiv* **217232** (2018).

31. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).

32. Cobey, S. & Lipsitch, M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Sci.* **335**, 1376–1380 (2012).

33. Lehtinen, S. *et al.* Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proc. Natl. Acad. Sci.* **114**, 1075–1080 (2017). URL http://www.pnas.org/content/114/5/1075. DOI 10.1073/pnas.1617849114.

34. Bell, B. G., Schellevis, F., Stobberingh, E., Goossens, H. & Pringle, M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. *BMC Infect. Dis.* **14**, 13 (2014). URL https://doi.org/10.1186/1471-2334-14-13. DOI 10.1186/1471-2334-14-13.

35. MacFadden, D. R. *et al.* A platform for monitoring regional antimicrobial resistance, using online data sources: Resistanceopen. *The J. Infect. Dis.* **214**, S393–S398 (2016).

36. Whitney, C. G. *et al.* Increasing prevalence of multidrug-resistant streptococcus pneumoniae in the united states. *New Engl. J. Medicine* **343**, 1917–1924 (2000).

37. McCormick, A. *et al.* Geographic diversity and temporal trends of antimicrobial resistance in *Streptococcus pneumoniae* in the united states. *Nat. Medicine* **9**, 424–431 (2003).

38. Fenoll, A. *et al.* Antimicrobial susceptibility and pneumococcal serotypes. *J. Antimicrob. Chemother.* **50**, 1–8 (2002).

39. Kimura, M. & Weiss, G. H. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genet.* **49**, 561–576 (1964).

40. Ball, F. Stochastic and deterministic models for sis epidemics among a population partitioned into households. *Math. Biosci.* **156**, 41–67 (1999).

41. Keeling, M. J. *Modeling infectious diseases in humans and animals* (Princeton University Press, Princeton, 2008).

42. Humplik, J., Hill, A. L. & Nowak, M. A. Evolutionary dynamics of infectious diseases in finite populations. *J. Theor. Biol.* **360**, 149–162 (2014). URL http://www.sciencedirect.com/science/article/pii/S0022519314003907. DOI 10.1016/j.jtbi.2014.06.039.

43. Lipsitch, M., Colijn, C., Cohen, T., Hanage, W. P. & Fraser, C. No coexistence for free: neutral null models for multistrain pathogens. *Epidemics* **1**, 2–13 (2009).

44. Bollobás, B. *Random Graphs* (Cambridge University Press, Cambridge, UK, 2001).

45. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nat.* **393**, 440–442 (1998).

46. Hay, S. I. & Snow, R. W. The Malaria Atlas Project: Developing Global Maps of Malaria Risk. *PLOS Medicine* **3**, e473 (2006). URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0030473. DOI 10.1371/journal.pmed.0030473.

47. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nat.* **496**, 504–507 (2013). URL http://www.nature.com/nature/journal/v496/n7446/full/nature12060.html. DOI 10.1038/nature12060.

48. Messina, J. P. *et al.* Mapping global environmental suitability for Zika virus. *eLife* **5**, e15272 (2016). URL https://doi.org/10.7554/eLife.15272. DOI 10.7554/eLife.15272.

49. Kraemer, M. U. G. *et al.* Progress and Challenges in Infectious Disease Cartography. *Trends Parasitol.* **32**, 19–29 (2016). URL http://www.sciencedirect.com/science/article/pii/S147149221500207X. DOI 10.1016/j.pt.2015.09.006.

50. Tatem, A. J. & Smith, D. L. International population movements and regional Plasmodium falciparum malaria elimination strategies. *Proc. Natl. Acad. Sci.* **107**, 12222–12227 (2010). URL http://www.pnas.org/content/107/27/12222. DOI 10.1073/pnas.1002971107.

51. Wesolowski, A. *et al.* Quantifying the impact of human mobility on malaria. *Sci.* **338**, 267–270 (2012). DOI 10.1126/science.1223467.

52. Carrasco-Escobar, G. *et al.* Micro-epidemiology and spatial heterogeneity of p. vivax parasitaemia in riverine communities of the peruvian amazon: A multilevel analysis. *Sci. Reports* **7**, 8082 (2017).

53. Kang, S. Y. *et al.* Heterogeneous exposure and hotspots for malaria vectors at three study sites in uganda. *bioRxiv* (2018).

54. Olesen, S. W. & Grad, Y. H. Racial/Ethnic Disparities in Antimicrobial Drug Use, United States, 2014–2015. *Emerg. Infect. Dis.* **24** (2018). URL https://wwwnc.cdc.gov/eid/article/24/11/18-0762_article. DOI 10.3201/eid2411.180762.

55. Olesen, S. W. *et al.* The distribution of antibiotic use and its association with antibiotic resistance. *eLife* **7**, e39435 (2018). URL https://doi.org/10.7554/eLife.39435. DOI 10.7554/eLife.39435.

56. Salathé, M. *et al.* A high-resolution human contact network for infectious disease transmission. *Proc. Natl. Acad. Sci.* 201009094 (2010). URL http://www.pnas.org/content/early/2010/12/08/1009094108. DOI 10.1073/pnas.1009094108.

57. Vanhems, P. *et al.* Estimating Potential Infection Transmission Routes in Hospital Wards Using Wearable Proximity Sensors. *PLoS ONE* **8**, e73970 (2013). URL http://dx.doi.org/10.1371/journal.pone.0073970. DOI 10.1371/journal.pone.0073970.

58. Voirin, N. *et al.* Combining High-Resolution Contact Data with Virological Data to Investigate Influenza Transmission in a Tertiary Care Hospital. *Infect. Control. & Hosp. Epidemiol.* **36**, 254–260 (2015). URL http://journals.cambridge.org/article_S0899823X14000531. DOI 10.1017/ice.2014.53.

59. Wesolowski, A. *et al.* The Use of Census Migration Data to Approximate Human Movement Patterns across Temporal Scales. *PLOS ONE* **8**, e52971 (2013). URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0052971. DOI 10.1371/journal.pone.0052971.

60. Tizzoni, M. *et al.* On the Use of Human Mobility Proxies for Modeling Epidemics. *PLOS Comput. Biol.* **10**, e1003716 (2014). URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003716. DOI 10.1371/journal.pcbi.1003716.

61. Wesolowski, A., O'Meara, W. P., Eagle, N., Tatem, A. J. & Buckee, C. O. Evaluating Spatial Interaction Models for Regional Mobility in Sub-Saharan Africa. *PLOS Comput. Biol.* **11**, e1004267 (2015). URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004267. DOI 10.1371/journal.pcbi.1004267.

62. Wesolowski, A., Buckee, C. O., Engø-Monsen, K. & Metcalf, C. J. E. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. *The J. Infect. Dis.* **214**, S414–S420 (2016). URL https://academic.oup.com/jid/article/214/suppl_4/S414/2527905. DOI 10.1093/infdis/jiw273.

63. Brockmann, D. & Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Sci.* **342**, 1337–1342 (2013). URL http://science.sciencemag.org/content/342/6164/1337. DOI 10.1126/science.1245200.

64. Nunes, M. R. T. *et al.* Air Travel Is Associated with Intracontinental Spread of Dengue Virus Serotypes 1–3 in Brazil. *PLOS Neglected Trop. Dis.* **8**, e2769 (2014). URL https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0002769. DOI 10.1371/journal.pntd.0002769.

65. Pybus, O. G. *et al.* Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proc. Natl. Acad. Sci.* **109**, 15066–15071 (2012). URL http://www.pnas.org/content/109/37/15066. DOI 10.1073/pnas.1206598109.

66. Snitkin, E. S. *et al.* Tracking a Hospital Outbreak of Carbapenem-Resistant Klebsiella pneumoniae with Whole-Genome Sequencing. *Sci. Transl. Medicine* **4**, 148ra116–148ra116 (2012). URL http://stm.sciencemag.org/content/4/148/148ra116. DOI 10.1126/scitranslmed.3004129.

67. Frost, S. D. W. *et al.* Eight challenges in phylodynamic inference. *Epidemics* **10**, 88–92 (2015). URL http://www.sciencedirect.com/science/article/pii/S1755436514000437. DOI 10.1016/j.epidem.2014.09.001.

68. Taylor, A. R. *et al.* Quantifying connectivity between local Plasmodium falciparum malaria parasite populations using identity by descent. *PLOS Genet.* **13**, e1007065 (2017). URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1007065. DOI 10.1371/journal.pgen.1007065.

69. Dudas, G. *et al.* Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nat.* **544**, 309–315 (2017). URL https://www.nature.com/articles/nature22040. DOI 10.1038/nature22040.

70. Fisher, R. A. The wave of advance of advantageous genes. *Annals Eugen.* **7**, 355–369. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1937.tb02153.x. DOI 10.1111/j.1469-1809.1937.tb02153.x. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-1809.1937.tb02153.x.

71. Kimura, M. & Weiss, G. H. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genet.* **49**, 561–576 (1964). URL http://www.genetics.org/content/49/4/561. http://www.genetics.org/content/49/4/561.full.pdf.

72. Levins, R. Theory of fitness in a heterogeneous environment. i. the fitness set and adaptive function. *The Am. Nat.* **96**, 361–373 (1962). URL http://www.jstor.org/stable/2458725.

73. Holt, R. D. Spatial heterogeneity, indirect interactions, and the coexistence of prey species. *The Am. Nat.* **124**, 377–406 (1984). URL http://www.jstor.org/stable/2461465.

74. Gavrilets, S. & Gibson, N. Fixation probabilities in a spatially heterogeneous environment. *Popul. Ecol.* **44**, 51–58 (2002).

75. Whitlock, M. & Gomulkiewicz, R. Probability of fixation in a heterogeneous environment. *Genet.* **171**, 1407–1417 (2005).

76. Lieberman, E., Hauert, C. & Nowak, M. A. Evolutionary dynamics on graphs. *Nat.* **433**, 312–316 (2005).

77. Manem, V., Kaveh, K., Kohandel, M. & Sivaloganathan, S. Modeling invasion dynamics with spatial random-fitness due to micro-environment. *PLoS One* **10**, e0140234 (2015).

78. Kaveh, K., McAvoy, A. & Nowak, M. The effect of spatial fitness heterogeneity on fixation probability. *arXiv* 1709.03031 (2017).

79. Leventhal, G. E., Hill, A. L., Nowak, M. A. & Bonhoeffer, S. Evolution and emergence of infectious diseases in theoretical and real-world networks. *Nat. Commun.* **6** (2015). URL http://www.nature.com/ncomms/2015/150116/ncomms7101/full/ncomms7101.html. DOI 10.1038/ncomms7101.

80. Heffernan, J., Smith, R. & Wahl, L. Perspectives on the basic reproductive ratio. *J. R. Soc. Interface* **2**, 281–293 (2005).

81. Kim, J. H. & Vu, V. H. Generating random regular graphs. *Comb.* **26**, 683–708 (2006).

82. Pundak, G. Random regular graph generator. *MATLAB Cent. File Exch.* (2010). URL https://www.mathworks.com/matlabcentral/fileexchange/29786-random-regular-generator.

83. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. Royal Stat. Soc. Ser. B (Methodological)* **58**, 267–288 (1996).

Supporting Information

Analytic results for well-mixed single deme with treatment

The dynamics of competition between drug-sensitive and drug-resistant strains of infection in an infinite, well-mixed population where a fraction f of individuals receive treatment if infected is described by Eqns. 1. This system of equations tracks the *proportion* of the population in each treatment and infection category (Infected with the wild type and untreated (w_u), infected with the wild type and treated (w_t), infected with the resistant strain and untreated (r_u), and infected with the resistant strain and treated (r_t)). These equations were developed by first writing down the dynamics of the *number* of individuals in each state, including uninfected (susceptible):

$$\begin{aligned}
\dot{S}_u &= -k\left(W_u + (1-\varepsilon)W_t + (1-c)(R_u+R_t)\right)S_u + g(W_u+R_u),\\
\dot{W}_u &= k\left(W_u + (1-\varepsilon)W_t\right)S_u - gW_u,\\
\dot{R}_u &= k(1-c)(R_u+R_t)S_u - gR_u,\\
\dot{S}_t &= -k\left(W_u + (1-\varepsilon)W_t + (1-c)(R_u+R_t)\right)S_t + g(W_t+R_t),\\
\dot{W}_t &= k(W_u + (1-\varepsilon)W_t)S_t - gW_t,\\
\dot{R}_t &= k(1-c)(R_u+R_t)S_t - gR_t.
\end{aligned} \tag{S.1}$$

The transmission rate k represents the rate at which each infected individual contacts and infects each susceptible individual.

Because the total population size N as well as the fraction of individuals who will (or will not) receive treatment if infected are constant, we can remove the variables for susceptible individuals using $S_u = N(1-f) - W_u - R_u$ and $S_t = Nf - W_t - R_t$. To further reduce the variables in this system and remove any dependence on total population size, we let $\kappa = kN$ and $x_i = X_i/N$ (the fraction of individuals instead of the total number, for $X = W, R$ and $i = u, t$), which then leads to Eqns. 1. By keeping $\kappa = kN$ as the constant parameter, we are implicitly assuming that the total rate at which a single individual can contact and infect others when the population is completely susceptible (kN) is independent of the population size. This is a reasonable assumption for many infections which require some type of active contact, and here allows us to reduce the number of parameters required. None of our results depend on this scaling, and the effect of changing population size is equivalent to changing κ in this deterministic model.

There are three possible equilibria of Eqns. (1). In each of these equilibria, there is no possible coexistence of infections. These equilibria are as follows:

1. No infection survives

$$w_u = r_u = w_t = r_t = 0 \tag{S.2}$$

This uninfected equilibrium is only stable when

$$\begin{aligned}
R_0^w &= \frac{\beta(1-f\varepsilon)}{g} < 1\\
R_0^r &= \frac{\beta(1-c)}{g} < 1
\end{aligned} \tag{S.3}$$

2. Only wild-type (drug-sensitive) strain survives

$$\begin{aligned}
w_u &= (1-f)\left(1 - \frac{1}{R_0^w}\right)\\
r_u &= 0\\
w_t &= f\left(1 - \frac{1}{R_0^w}\right)\\
r_t &= 0
\end{aligned} \tag{S.4}$$

This infected equilibrium is stable when

$$\begin{aligned}
R_0^w &> 1\\
R_0^r &< R_0^w
\end{aligned} \tag{S.5}$$

where the later condition implies $c > f\varepsilon$.

3. Only drug-resistant strain survives

$$
\begin{aligned}
w_u &= 0 \\
r_u &= (1-f)\left(1-\frac{1}{R_0^r}\right) \\
w_t &= 0 \\
r_t &= f\left(1-\frac{1}{R_0^r}\right)
\end{aligned}
\tag{S.6}
$$

This infected equilibrium is stable when

$$
\begin{aligned}
R_0^r &> 1 \\
R_0^w &< R_0^r
\end{aligned}
\tag{S.7}
$$

where the former condition implies $c < f\varepsilon$.

Analytic results for two connected subpopulations

In an effort to be more general than in the main text, we consider the case of a population of total size N divided into two demes, with a fraction f of the total population is contained in the drug-treated deme. The system describing the *number* of individuals in each state can then be written as:

$$
\begin{aligned}
\dot{S}_u &= -(kW_u + b(1-\varepsilon)W_t + (1-c)(kR_u + bR_t))S_u + g(W_u + R_u), \\
\dot{W}_u &= (kW_u + b(1-\varepsilon)W_t)S_u - gW_u, \\
\dot{R}_u &= (1-c)(kR_u + bR_t)S_u - gR_u, \\
\dot{S}_t &= -S_t(bW_u + k(1-\varepsilon)W_t + (1-c)(bR_u + kR_t)) + g(W_t + R_t), \\
\dot{W}_t &= (bW_u + k(1-\varepsilon)W_t)S_t - gW_t, \\
\dot{R}_t &= (1-c)(bR_u + kR_t)S_t - gR_t.
\end{aligned}
\tag{S.8}
$$

Since the size of each deme is fixed, we can remove the variables for susceptible individuals using $S_u = N(1-f) - W_u - R_u$ and $S_t = Nf - W_t - R_t$. To create a nondimensional version of the system, we instead track the *fraction* of individuals in each deme who have a particular infection status, so $w_u = W_u/(N(1-f))$, $R_u = R_u/(N(1-f))$, $w_t = W_t/(Nf)$, $R_t = R_t/(Nf)$, and $\kappa = kN$ and $\beta = bN$, which gives the altered system:

$$
\begin{aligned}
\dot{w_u} &= (\kappa(1-f)w_u + \beta f(1-\varepsilon)w_t)(1-w_u-r_u) - gw_u, \\
\dot{r_u} &= (1-c)(\kappa(1-f)r_u + \beta f r_t)(1-w_u-r_u) - gr_u, \\
\dot{w_t} &= (\beta(1-f)w_u + \kappa f(1-\varepsilon)w_t)(1-w_t-r_t) - gw_t, \\
\dot{r_t} &= (1-c)(\beta(1-f)r_u + \kappa f r_t)(1-w_t-r_t) - gr_t.
\end{aligned}
\tag{S.9}
$$

Using the next-generation technique[80], we arrive at the following values for the basic reproductive ratios,

$$
\begin{aligned}
R_0^w &= \frac{1}{2}\frac{\kappa}{g}\left((1-f\varepsilon) + \sqrt{(1-f\varepsilon)^2 - 4\left(1-\frac{\beta^2}{\kappa^2}\right)f(1-f)(1-\varepsilon)}\right), \\
R_0^r &= \frac{1}{2}\frac{\kappa}{g}(1-c)\left(1 + \sqrt{1 - 4\left(1-\frac{\beta^2}{\kappa^2}\right)f(1-f)}\right).
\end{aligned}
\tag{S.10}
$$

Equations (S.9) and (S.10) reduce to Eq. (2) and

In general, this system has nine equilibria. However, only one of these is stable and physically realizable (e.g. having between 0% and 100% of individuals infected) for any given parameter set. These are the values that are plotted in the main text. These outcomes fall into one of two qualitative categories. Either one strain drives the other to competitive extinction (true in Fig. 3 for $c > 0.4$ and $c < 0.06$, for instance) or both strains exist in both demes simultaneously. The boundaries separating this region from the competitive-exclusion region are $R_0^r = R_0^w$, the lower bound when viewed as a function of the cost of resistance c, and the curve which solves $\beta(1-f)w_u^* + \kappa f(1-\varepsilon)w_t^* = \beta(1-f)(1-c)r_u^* + \kappa f(1-c)r_t^*$, where $[w_u^*, w_t^*, r_u^*, r_t^*]$ denote the equilibrium values. This second curve, in words, is where the drug-treated population is infected at the same rate by both drug-resistant and drug-sensitive strains.

For the parameters chosen here and used throughout the paper, it is impossible to have $R_0^w < 1$. For sufficiently high c, we can have $R_0^r < 1$. However, this line in parameter space lies deep within the region where the population is infected with the wild-type strain only.

Scaling and nondimensionalization of the multi-deme model

The dynamics of infection in a network of interconnected demes can be described by the following system of equations

$$\begin{aligned}
\dot{S}_i &= -\left(k(1-\varepsilon T_i)W_i + b\sum_j \Delta_{ji}(1-\varepsilon T_j)W_j + k(1-c)R_i + b(1-c)\sum_j \Delta_{ji}R_j\right)S_i + g(W_i+R_i),\\
\dot{W}_i &= k(1-\varepsilon T_i)W_iS_i + bS_i\sum_j \Delta_{ji}(1-\varepsilon T_j)W_j - gW_i,\\
\dot{R}_i &= k(1-c)R_iS_i + b(1-c)S_i\sum_j \Delta_{ji}R_j - gR_i,
\end{aligned} \tag{S.11}$$

where X_i represents the total number of individuals in deme i infected with strain X (or susceptible, S) and D_i is the total number of individuals in deme i. Assuming all demes are of the same fixed size D allows us to remove the equations for the susceptible individuals (since $D_i = S_i + W_i + R_i$) and scale the system on D and remove it as a parameter. Writing $x_i = X_i/D$, $\kappa = kD$ and $\beta = bD$ gives us the reduced system presented in Eqs (3).

The apparent absence of multiple equilibria

Although the generic system of equations modeled by Eqns. (3) admits multiple equilibria, we actually never observed any dynamic phenomena other than a steady convergence to a single stable fixed point, even when initial conditions were wildly different from that considered in the main text. It is not uncommon for infection models to admit equilibria which are physically unrealizable, i.e., taking values outside the unit simplex — it may therefore be the case that any other stable equilibria are simply unphysical.

However, to comprehensively demonstrate this fact — either theoretically or numerically — is outside the scope of the current work. To prove such a result theoretically would necessarily draw on ideas of nonlinear algebra or algebraic geometry, such as intersection theory, which are outside the present scope; to demonstrate numerically that for all parameter sets and graph structures there is only one stable equilibrium is computationally infeasible, even on graphs of a given small size.

To illustrate the idea that there seems to be only one stable equilibrium per system, we fixed an exact distribution of treatment with $\rho = 0.35$ to be used on every random regular graph used in the main text. This means that, since all nodes in all graphs are labelled (and the random regular graphs in the main text have twenty demes each), seven labelled demes (here 1, 3, 8, 11, 15, 19, 20) were drug-treated irrespective of graph structure. We then performed, for all 30 graphs, 100 trials with uniformly random initial conditions over every deme, i.e., every deme can start with any possible arrangement of wild-type and resistant infectives (so long as the sum of the two does not exceed 1). We then compared the equilibria seen with the equilibria found for the initial condition used in the main body of the text.

There are many possible metrics which would illustrate the relative closeness of all the equilibria seen. In Fig. S12 we have elected to plot the maximum matrix norm (the 2-norm) of the difference between the equilibrium seen in any of the 100 trials and the equilibrium from the main text. That is to say, all of the trial equilibria differ in norm from the main-text equilibrium by an amount not more than that plotted in the figure for a given graph structure. The parameters used were our default $(\kappa, \beta, g, \varepsilon, c) = (0.25, 0.05, 0.1, 0.9, 0.2)$.

In all of these cases, the maximum norm deviation is extremely small. This could be due to two possible reasons. The first is due simply to aggregated numerical error from the many steps involved of a) calculating the equilibrium (which involves a threshold of stopping when the time-derivatives in all demes fall below a certain threshold, which can allow for some very small freedom in what an "equilibrium" is in a given trial), b) calculating the matrix norm (which amplifies these small errors). The second explanation is that while multiple stable equilibria may exist, they are extremely close together, with almost clinically-unmeasurable differences in the rates of wild-type and resistant infections within a given deme in a given community.

Model with effects of treatment on recovery rates

To demonstrate the robustness of the results described in the main text, we explored a second model in which the drug, rather than inhibiting the spread of the infection, instead speeds up recovery. The effect of treatment is to increase the rate of recovery from g to $g+\tau$. Here, τ represents the idea that an infected individual will seek a clinician, obtain a prescription for drug, and clear the infection as a result of drug at times represented by Poisson distributions, the average of which is τ. While in a mathematical sense there is no bound on τ in the way there is a bound on ε in the original model (where $\varepsilon < 1$), there is a realistic range for τ. For instance, if $(g+\tau) = 1$, then it would take the average infected individual only one day to be diagnosed, prescribed, and cured – this is perhaps an unrealistic situation to model. We therefore usually used a value of τ between 0.1 and 0.3.

This model matches qualitatively with that discussed in the main text, in that in reproduces the spatial motifs in the same way. The two models are also quantitatively quite similar for reasonably-akin parameter sets. Here we reproduce the main results for this alternative model.

A single well-mixed deme

Here, the system can be written as

$$\begin{aligned}
\dot{w_u} &= \kappa(w_u+w_t)((1-f)-w_u-r_u)-gw_u,\\
\dot{r_u} &= \kappa(1-c)(r_u+r_t)((1-f)-w_u-r_u)-gr_u,\\
\dot{w_t} &= \kappa(w_u+w_t)(f-w_t-r_t)-(g+\tau)w_t,\\
\dot{r_t} &= \kappa(1-c)(r_u+r_t)(f-w_t-r_t)-gr_t.
\end{aligned} \tag{S.12}$$

Here, R_0 for the resistant strain is unchanged from the main text. For the wild-type strain,

$$R_0^w = \frac{\kappa}{g+f\tau} \tag{S.13}$$

For the purpose of comparing this model, the value of τ for which R_0^w in the main-text model (with our baseline $\varepsilon = 0.9$, $f = 0.5$) equals R_0^w in this alternative model is $\tau = 0.163$.

Two connected demes

In this model, the equivalent of Eqs S.9 can be written as

$$\begin{aligned}
\dot{w_u} &= (\kappa(1-f)w_u+\beta f w_t)(1-w_u-r_u)-gw_u,\\
\dot{r_u} &= (1-c)(\kappa(1-f)r_u+\beta f r_t)(1-w_u-r_u)-gr_u,\\
\dot{w_t} &= (\beta(1-f)w_u+\kappa f w_t)(1-w_t-r_t)-(g+\tau)w_t,\\
\dot{r_t} &= (1-c)(\beta(1-f)r_u+\kappa f r_t)(1-w_t-r_t)-gr_t.
\end{aligned} \tag{S.14}$$

Using the next-generation technique, we were able to derive the values of R_0 in a two-deme system for the alternative model, as an analog to Eq. S.10:

$$\begin{aligned}
R_0^{WT} &= \frac{1}{2g(g+\tau)}(g\kappa-(g-(1-f)\kappa)\tau\\
&\quad+\sqrt{g^2(\kappa^2+4(1-f)f(\beta-\kappa)(\beta+\kappa)+2g(2(1-f)f\beta^2+(1-2f)g\kappa-(1-f)(-1+2f)\kappa^2)\tau+(g+\kappa-f\kappa^2)\tau^2)},\\
R_0^r &= \frac{1}{2}\frac{\kappa}{g}(1-c)\left(1+\sqrt{1-4\left(1-\frac{\beta^2}{\kappa^2}\right)f(1-f)}\right).
\end{aligned} \tag{S.15}$$

The expressions are slightly more complicated, but the curves generated are very similar to the main text. Indeed, Fig. S9 shows the analog of Fig. 3, and the two are very similar upon visual inspection. As in the main text, here the boundaries of the coexistence region are the curves $R_0^w = R_0^r$ (lower boundary) and the implicit curve where more wild-type infectives than resistant infectives are produced per unit time in the drug-treated deme (upper boundary).

Scaling and nondimensionalization of the multi-deme model

Here, the model can be summarized as a set of ODEs akin to Eq. (1) for each deme:

$$\begin{aligned}
\dot{W_i} &= kW_i(D-W_i-R_i)-(g+T_i\tau)W_i+b(D-W_i-R_i)\sum_j \Delta_{ji}W_j,\\
\dot{R_i} &= k(1-c)R_i(D-W_i-R_i)-gR_i+b(1-c)(D-W_i-R_i)\sum_j \Delta_{ji}R_j.
\end{aligned} \tag{S.16}$$

As in the main model, writing $w_i = W_i/D$, $r_i = R_i/D, \kappa = kD$ and $\beta = bD$ gives a nondimensionalized system,

$$\begin{aligned}
\dot{w_i} &= \kappa w_i(1-w_i-r_i)-(g+T_i\tau)w_i+\beta(1-w_i-r_i)\sum_j \Delta_{ji}w_j,\\
\dot{r_i} &= \kappa(1-c)r_i(1-w_i-r_i)-gr_i+\beta(1-c)(1-w_i-r_i)\sum_j \Delta_{ji}r_j.
\end{aligned} \tag{S.17}$$

Supplementary Figures

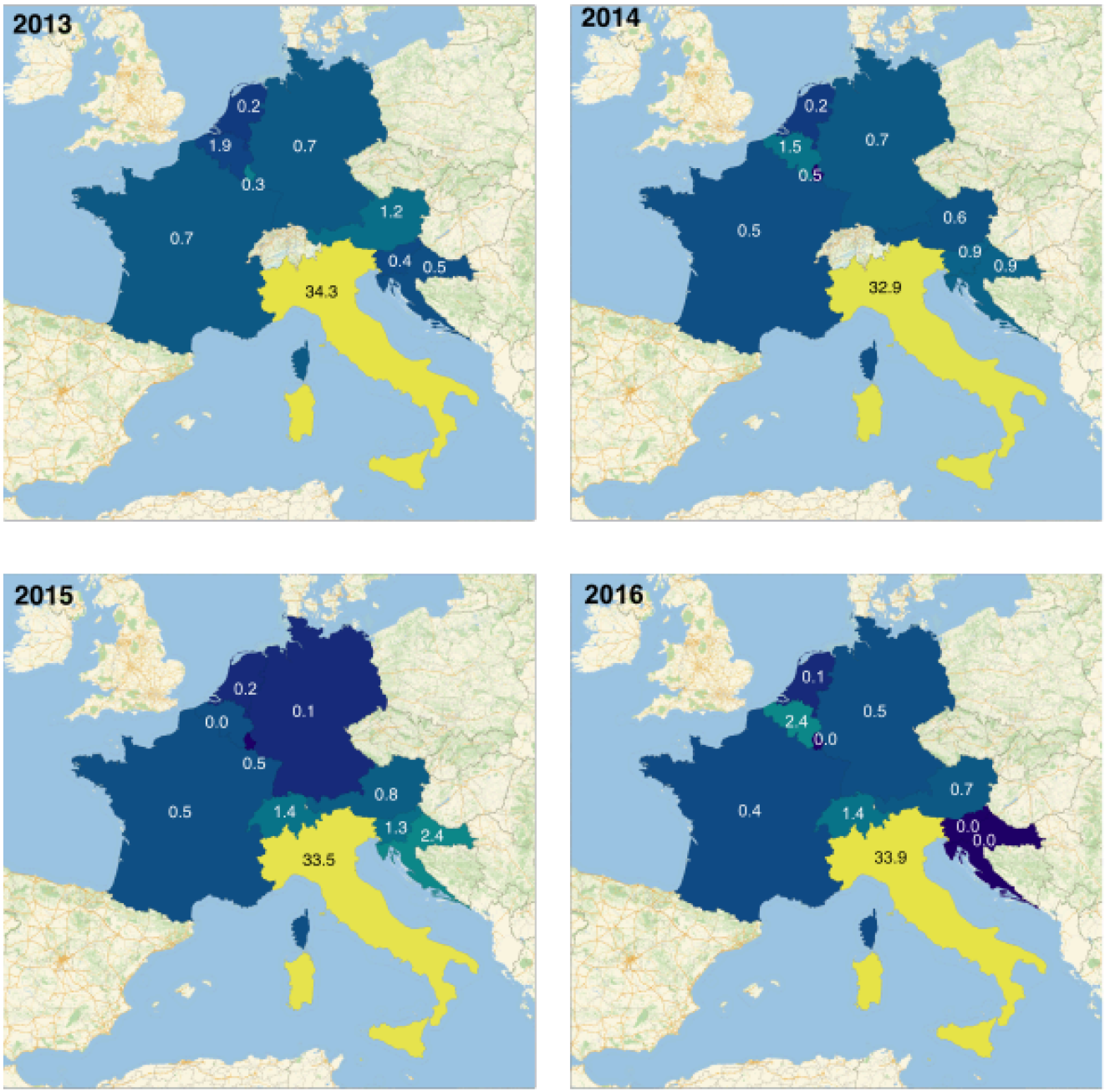

Figure S1. Stable spatial heterogeneity in resistance levels in Europe. Percent of *Klebsiella pneumoniae* isolates resistant to carbapenems in Austria, Belgium, Croatia, France, Germany, Italy, Luxembourg, and Slovenia from 2013-2016 as reported in the ECDC's European Antibiotic Resistance Surveillance Network[13] and Switzerland as reported by the Swiss Centre for Antibiotic Resistance from 2015-2016. Each country is labeled with the resistance level (%). The year-to-year deviation in Italy from the average value shown in Figure 1 is less than 1% for all years. For all other countries, the frequency of resistance never reaches even $1/10$ of the average value seen in Italy.

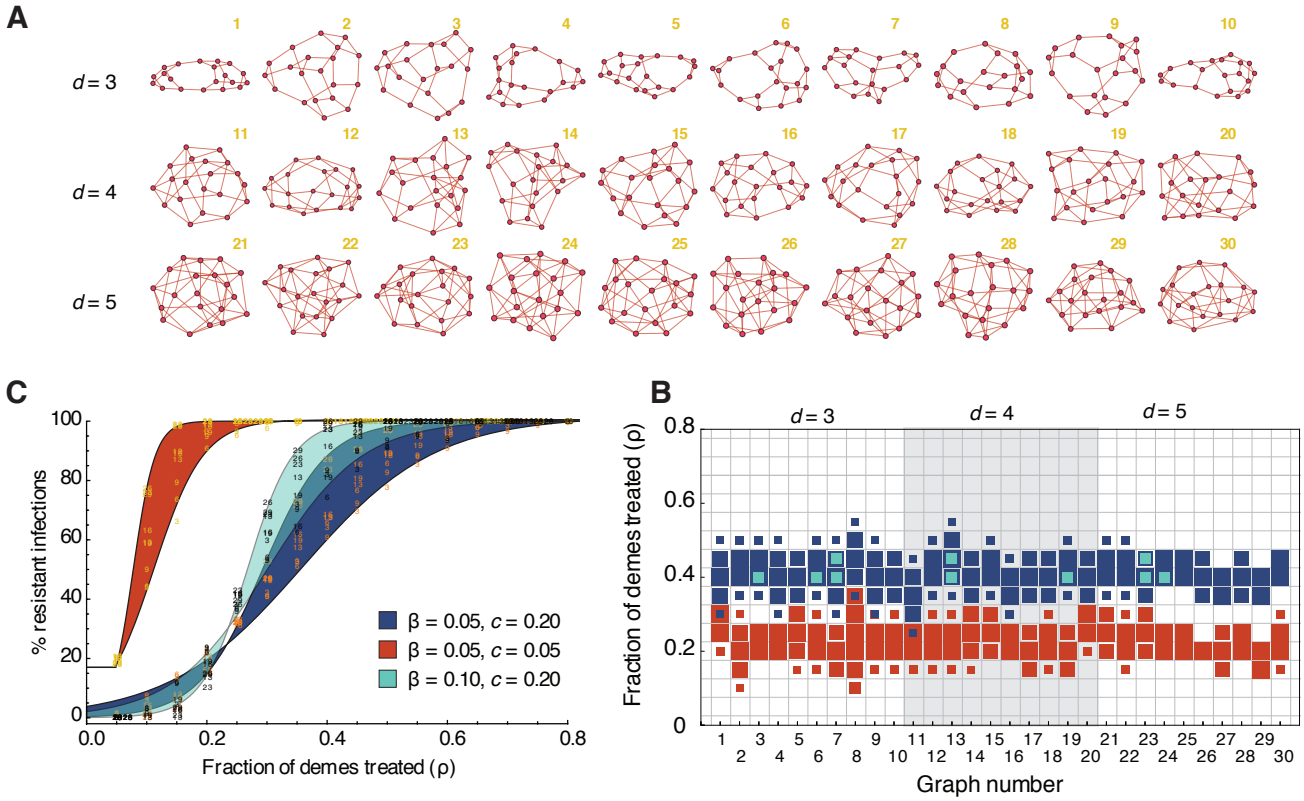

Figure S2. Dynamics of drug-resistant infections in populations consisting of networks of inter-connected demes with varying degree. A) Randomly generated population structures on which infection was simulated. Each node represents a deme (a well-mixed sub-population of individuals), and each edge indicates that infection can spread in either direction between those two demes. The total population consists of $M = 20$ demes and each deme has either d =3, 4, or 5 randomly connected neighbors. Ten example populations for each d were created and indexed for infection simulations. **B)** Fraction of infections that are resistant in the entire population (y-axis) versus fraction of demes treated, ρ (x-axis). Each color represents a different parameter set (blue background — baseline, red background — lower cost of resistance, teal background — more between-deme connectivity). Numbers show data points for individual populations indexed in (A). For each population, 100 simulations were run, each with a different spatial arrangment of treatment but perserving overall treatment levelρ. The colored envelope is created by shading between sigmoidal curves that were fit to the results for the $d = 3$ and $d = 5$ graphs. **C)** For each example population structure shown in (A) (x-axis) and each treatment level (ρ), the proportion of simulations that resulted in *robust* coexistence between drug-sensitive and drug-resistant strains is shown (by the colored area of the box). Robust coexistence was defined as at least 80% of demes supporting both strains at frequencies above 10%. All simulations used kinetic parameters $\kappa = 0.25$, $g = 0.1$, and $\varepsilon = 0.9$.

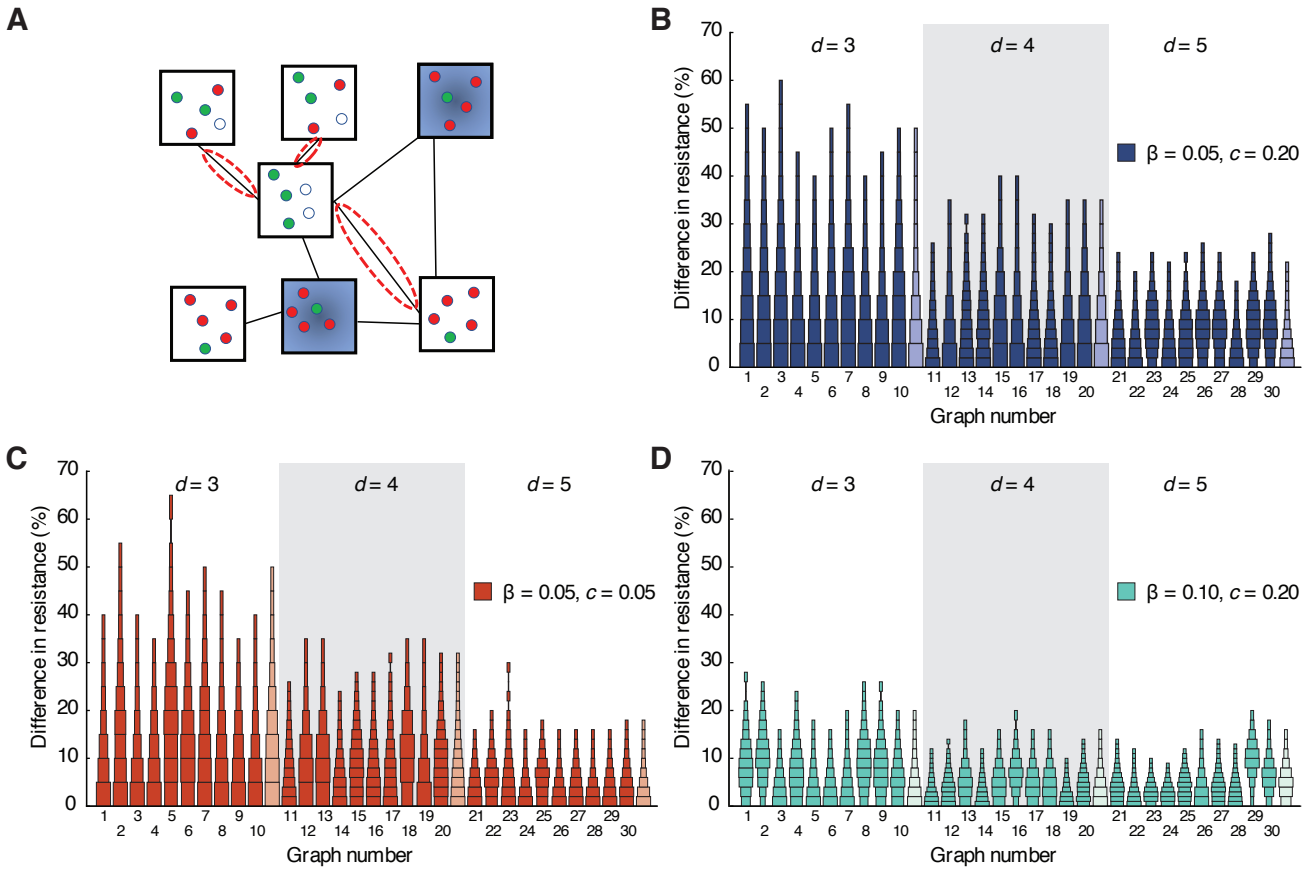

Figure S3. Comparing resistance levels between neighboring regions with identical treatment coverage, as a function of network degree. A) Differences in resistance levels (% of all infections that are with the drug resistant-strain) are measured between all pairs of directly-connected untreated demes. **B)-D)** Histograms showing the distribution of pairwise differences in resistance for a given population structure (indexed in Fig. S2A), pooling results from 100 simulations with different random allocations of treatment across demes. Lighter shaded histograms combine results from all population graphs with a given degree (d=3, 4, or 5). **B)** Baseline parameter values. **C)** Lower cost of resistance. **D)** Higher between-deme transmission rate. All simulations used kinetic parameters $\kappa = 0.25$, $g = 0.1$, and $\varepsilon = 0.9$ and an overall treatment fraction of $\rho = 0.3$.

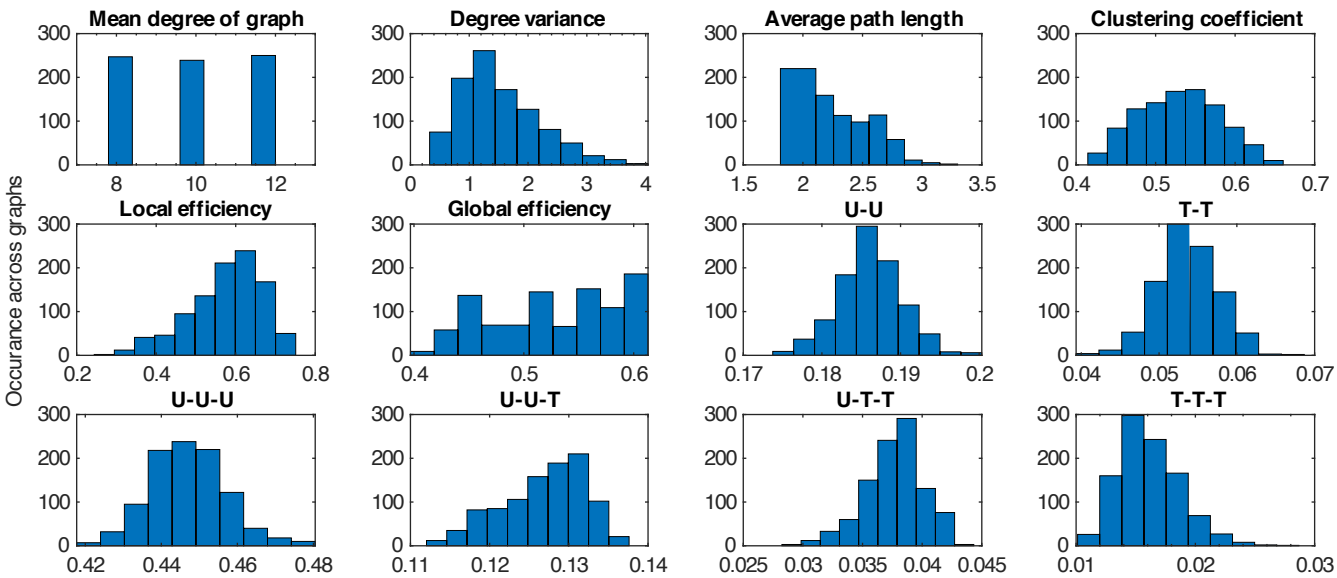

Figure S4. The distribution of graph-theoretic properties of simulated populations used in LASSO regression analysis. Populations were created as random networks of 50 demes using a variant of the Watts-Strogatz algorithm. Treatment was randomly assigned to a portion ρ of demes so that an overall desired fraction treated was achieved (here, $\rho = 0.24$). Distributions show properties for 1000 such networks. The first six properties are intrinsic to the network structure, while the latter six describe how treatment is allocated over the network. Quantities of the form X-Y give the proportion of all pairs of connected demes in which one deme has treatment status X and one has treatment status Y (U=untreated, T=treated). Quantities of the form X-Y-Z give the same information for triples of connected demes (order ignored). A description of the methods for creating the networks and for calculating each property is given in the **Methods**.

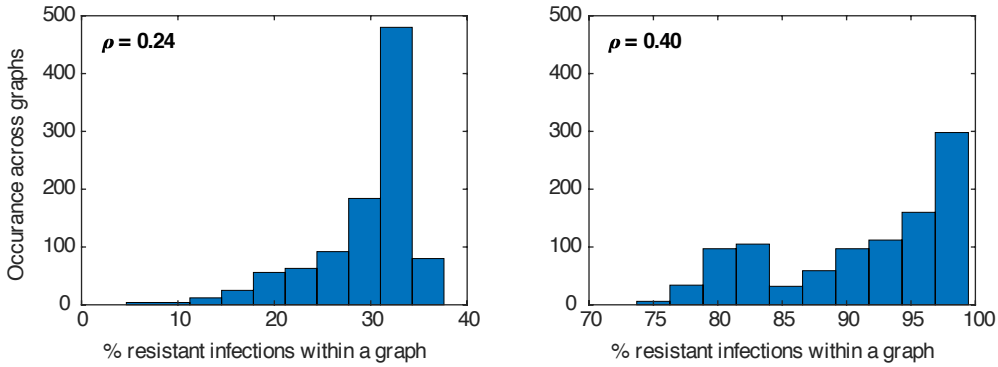

Figure S5. The distribution of the fraction of all infections which are drug-resistant in the simulated populations used in the LASSO regression analysis. Populations were created as random networks of 50 demes using a variant of the Watts-Strogatz algorithm. Treatment was randomly assigned to a portion ρ of demes so that an overall desired fraction treated was achieved (on the left $\rho = 0.24$, and on the right $\rho = 0.4$). Infection dynamics were simulated (Eq. (3)) until an equilibrium was reached, at which the proportion of all infections with the drug-resistant strain was recorded. Distributions show results for 1000 such simulations each with a unique random network and treatment allocation. Parameters used were $\kappa = 0.25$, $\beta = 0.05$, $g = 0.1$, $\varepsilon = 0.9$, $c = 0.2$.

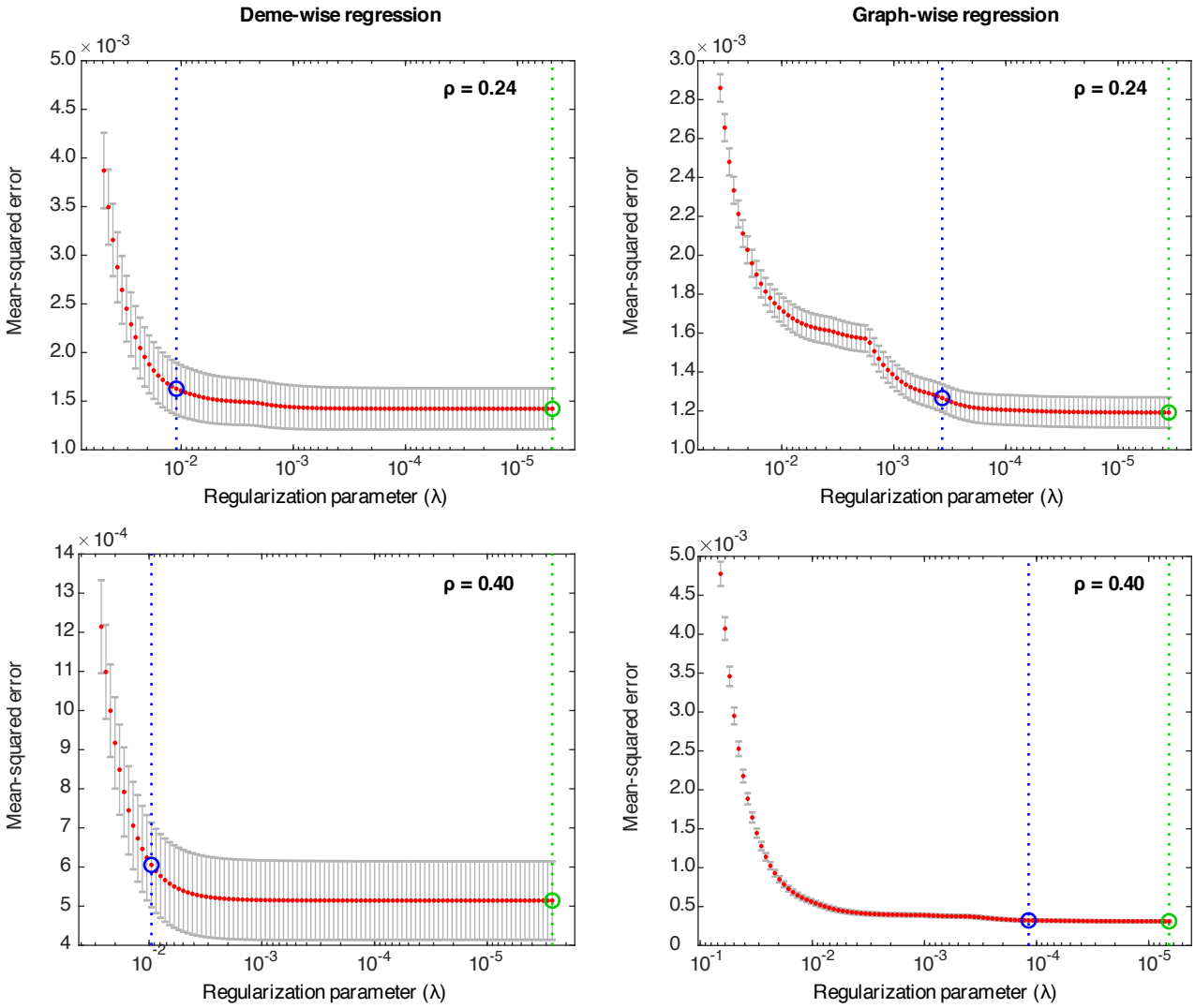

Figure S6. Five-fold cross-validation of LASSO regression models. Mean-squared error of model predictions vs sub-sample of 1/5 of data after model fitting to 4/5 of data, as a function of the internal parameter penalizing the L_1 norm for fitted coefficients (λ). Red dots show the mean-squared error of the model, with grey bars showing the standard error of the mean. The green circle/line indicates the value of λ that gives the minimum cross-validation error, whereas the blue circle/line locates the λ value where the mean-squared error is one standard error above that of the λ value with minimum cross-validation error.

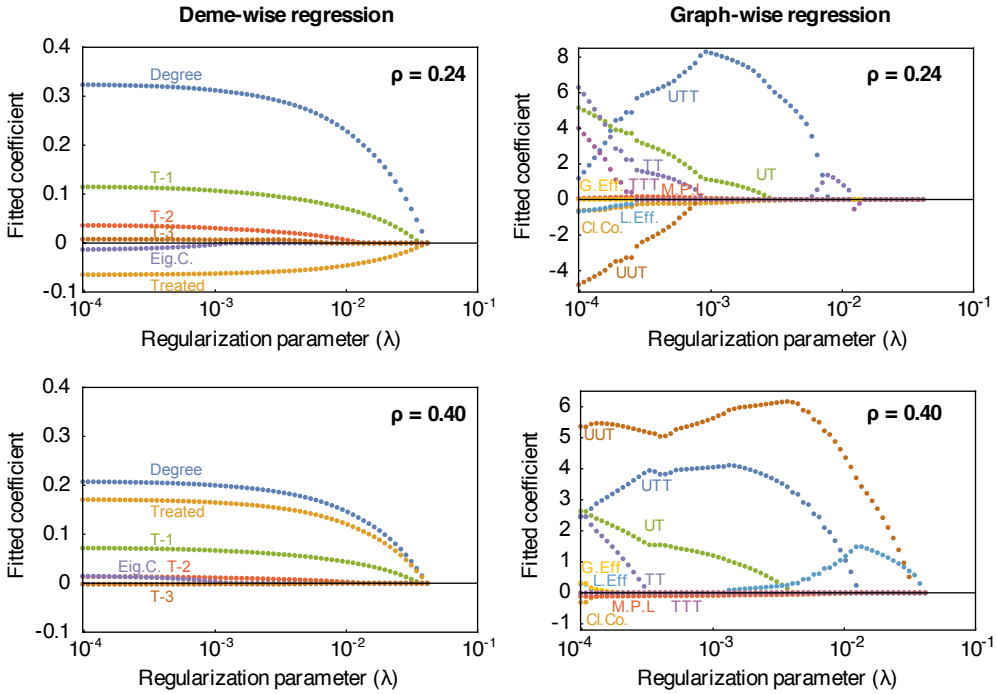

Figure S7. Results of LASSO regression to identify predictors of resistance. Plots show the regression coefficients obtained as a function of the internal constraint on L_1 norm of coefficients ("regularization parameter", λ). Each curve is a different predictor variable (described in Table 1). Analysis were separately run to predict deme-level resistance from deme-level properties (left side) or population-level resistance from graph-level properties (right side). Each analysis was conducted for two different levels of drug coverage in the population (either a fraction $\rho = 0.24$ or $\rho = 0.4$ of demes treated). The labels on the curves match the names of the predictors in Table 1.

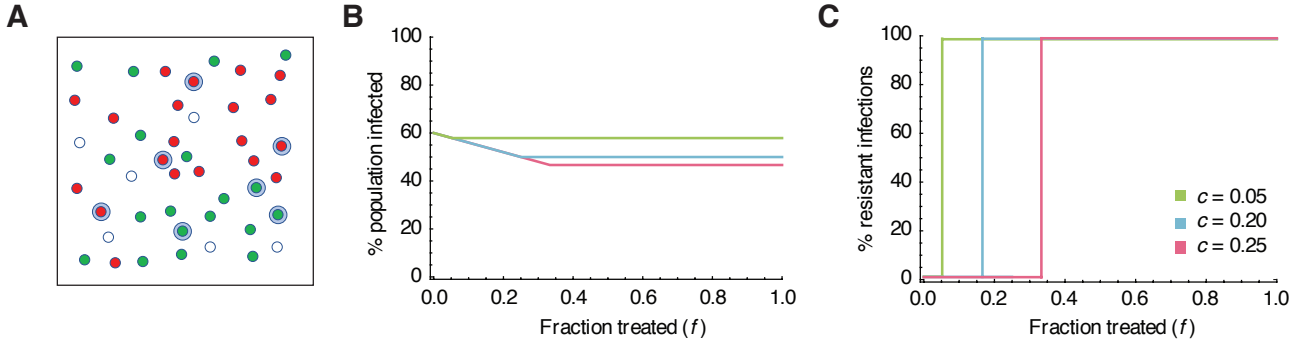

Figure S8. Analytical results on the equilibria found in one deme in the model where drug treatment increases the recovery rate from wild-type infection. A) The model in one deme. Individuals are either susceptible, infected with the wild-type strain, or infected with drug-resistant strain. A fraction f of individuals will be drug-treated if infected. **B)** The percentage of infected individuals (with either strain) as a function of the treated fraction f, for $\kappa = 0.25$, $g = 0.1$, $\tau = 0.1$ and the values of c shown in the legend in the rightmost panel. **C)** The percentage of infections which are drug-resistant as a function of the treated fraction f for the kinetic parameters described in (B). Since this percentage jumps from 0% to 100% at a critical value f^* depending on kinetic parameters, there is no coexistence in one deme in this model.

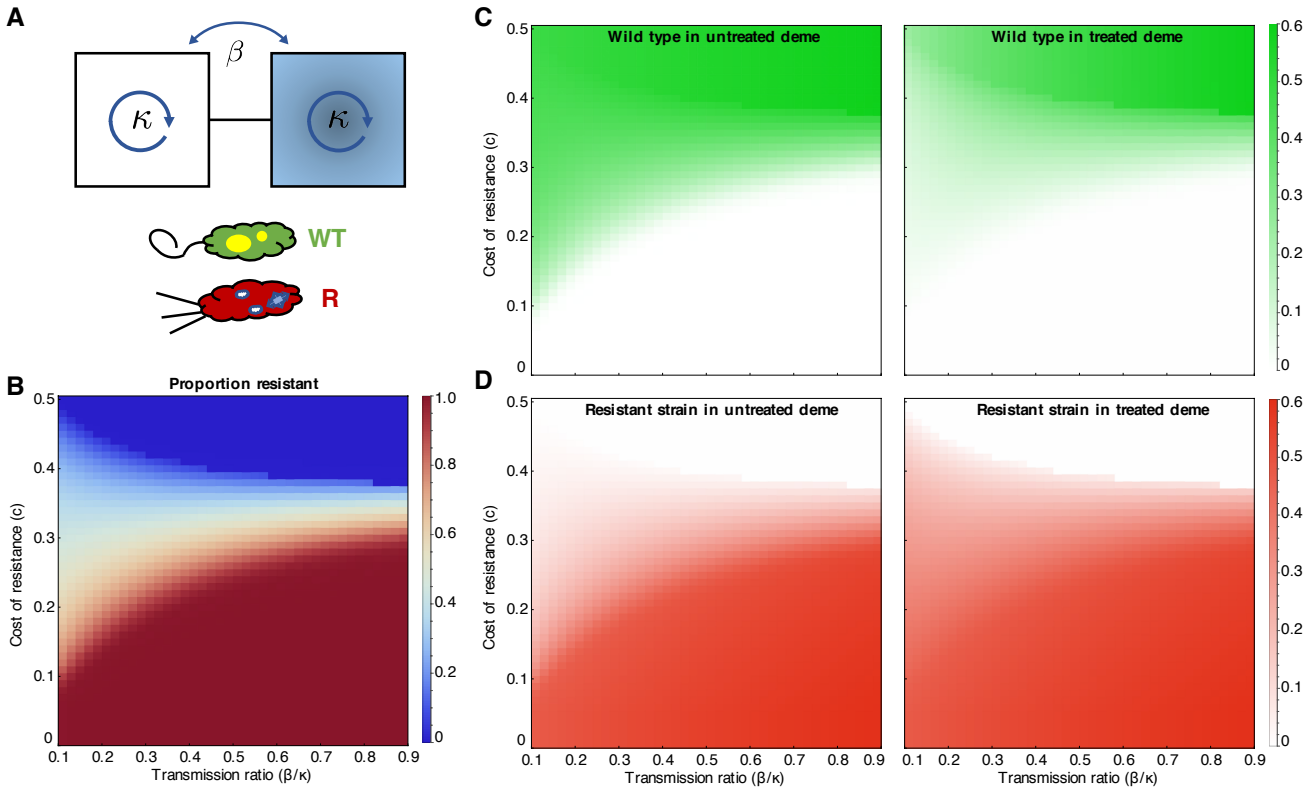

Figure S9. Competition between drug-sensitive and drug-resistant strains in two interconnected subpopulations, when treatment increases recovery rate. Competition between wild-type and resistant strains in two interconnected subpopulations. Each panel shows the infection level (shading) as a function of the level of connectivity between demes (β/κ) and the cost of resistance (c). **A)** Schematic of the two-demes (left-untreated, right-treated) and two strains (green-wild type, red-resistant) considered in the model. **B)** The overall frequency of resistant infections (fraction of all infections that are with drug-resistant strain). **C)** The fraction of individuals in each deme who are infected with the wild type strain. **D)** The fraction of individuals in each deme who are infected with the resistant strain. For these results, the kinetic parameters are $\kappa = 0.25$, $g = 0.1$, $\tau = 0.15$.

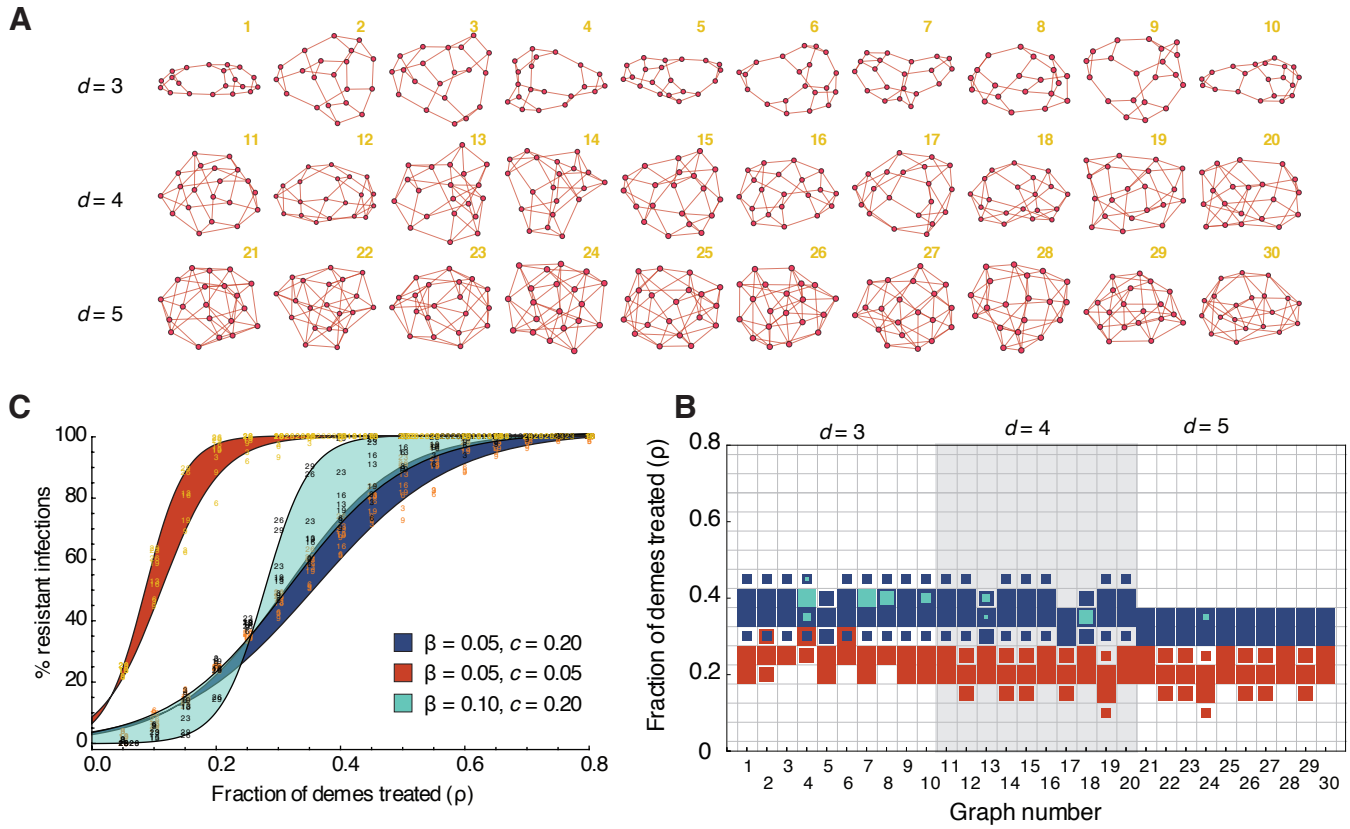

Figure S10. Dynamics of drug-resistant infections in populations consisting of networks of inter-connected demes, when treatment increases recovery rate. **A)** Randomly generated population structures on which infection was simulated. Each node represents a deme (a well-mixed sub-population of individuals), and each edge indicates that infection can spread in either direction between those two demes. The total population consists of $M = 20$ demes and each deme has either d =3, 4, or 5 randomly connected neighbors. Ten example populations for each d were created and indexed for infection simulations. **B)** Fraction of infections that are resistant in the entire population (y-axis) versus fraction of demes treated, ρ (x-axis). Each color represents a different parameter set (blue background — baseline, red background — lower cost of resistance, teal background — more between-deme connectivity). Numbers show data points for individual populations indexed in (A). The colored envelope is created by shading between sigmoidal curves that were fit to the results for the 3-regular and 5-regular graphs. **C)** For each example population structure shown in (A) (x-axis) and each treatment level (ρ), the proportion of simulations that resulted in *robust* coexistence between drug-sensitive and drug-resistant strains is shown (by the colored area of the box). Each of the 100 simulations for each population structure had a different random spatial arrangement of treatment, where with probability ρ individuals in a given deme would all receive treatment if infected. Robust coexistence was defined as at least 80% of demes supporting both strains at frequencies above 10%.

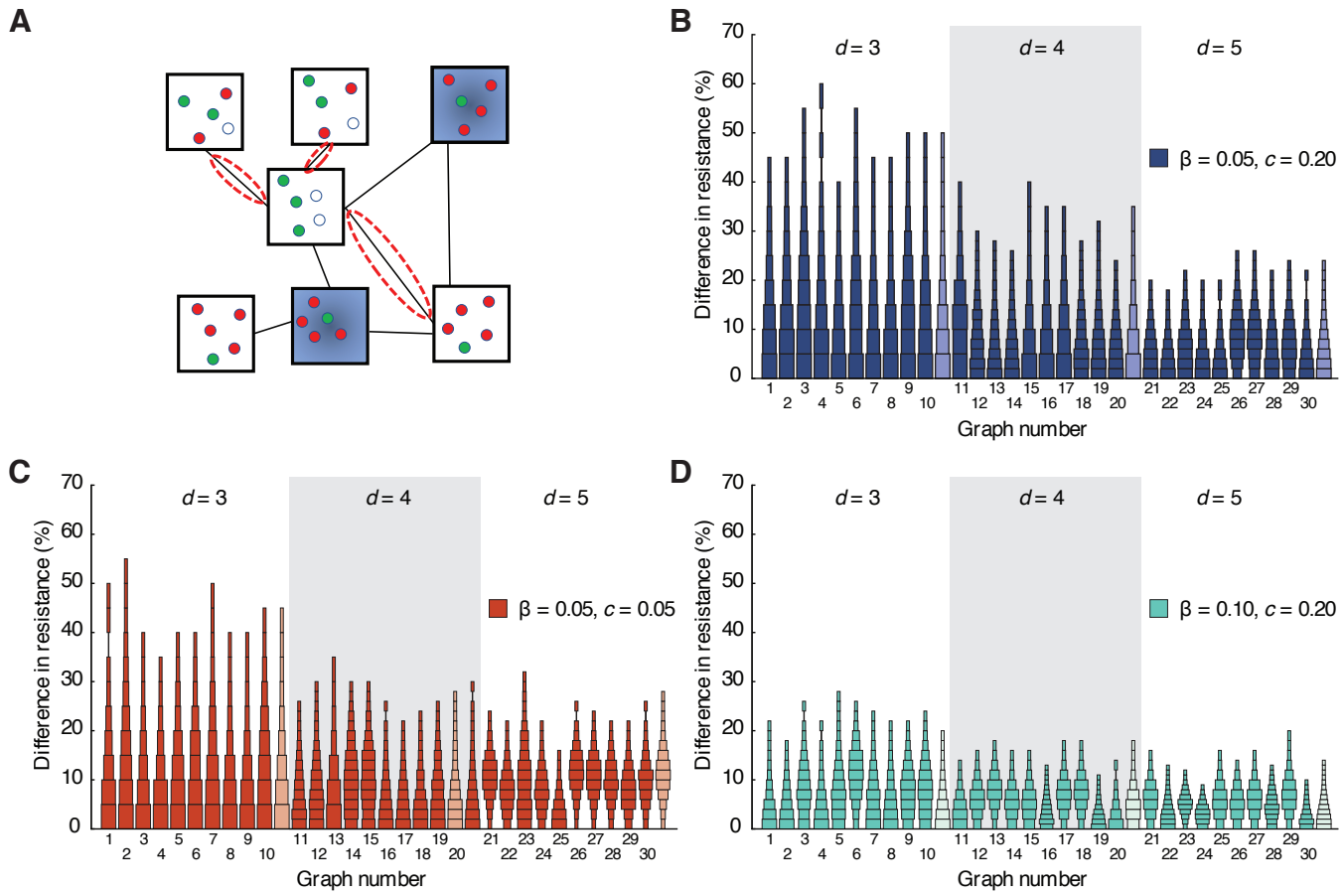

Figure S11. Comparing resistance levels between neighboring regions with identical treatment coverage, when treatment increases recovery rate. **A)** Pairwise differences in resistance are measured between all untreated demes sharing an edge. **B)-D)** The distribution of the differences in resistance across such pairwise edges in the population structures and for the kinetic parameters used in Fig. S2 after many realizations are collated. Numbers along the x axis indicate the population structure; bolded indicators **d3**, **d4**, and **d5** indicate the distributional average for all graphs of degree $d = 3$, 4, and 5, respectively. Globally, a fraction $\rho = 0.3$ of demes are drug-treated.

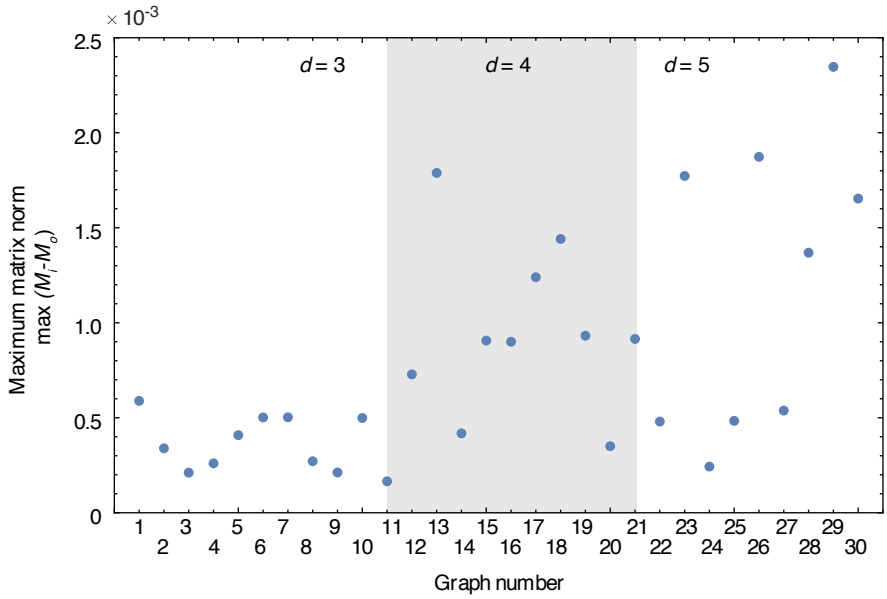

Figure S12. Difference between equilibrium values obtained for infection dynamics in multi-deme systems as initial conditions vary. For each of the population structures depicted in Fig. S2A, treatment was allocated randomly across demes with an overall proportion treated $\rho = 0.35$. For each graph-treatment allocation combination, 100 different sets of initial conditions (levels of resistant and sensitive infections in each deme) were chosen uniformly at randomly, and infection dynamics were run until an equilibrium was reached as described in the Methods. The maximum difference (in matrix 2-norm) was calculated between the equilibria seen in any of 100 trials and the equilibrium values used for results reported in the main text. Very low maximum norm values suggest there is a single stable equilibrium value for each parameter set.