Host population structure and treatment frequency maintain balancing selection on drug resistance

Sarah Cobey¹, Edward B. Baskerville¹, Caroline Colijn², William Hanage³, Christophe Fraser⁴ and Marc Lipsitch³

¹Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA
²Department of Mathematics, Imperial College London, London, UK
³Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA
⁴Oxford Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK

It is a truism that antimicrobial drugs select for resistance, but explaining pathogen- and population-specific variation in patterns of resistance remains an open problem. Like other common commensals, *Streptococcus pneumoniae* has demonstrated persistent coexistence of drug-sensitive and drug-resistant strains. Theoretically, this outcome is unlikely. We modelled the dynamics of competing strains of *S. pneumoniae* to investigate the impact of transmission dynamics and treatment-induced selective pressures on the probability of stable coexistence. We find that the outcome of competition is extremely sensitive to structure in the host population, although coexistence can arise from age-assortative transmission models with age-varying rates of antibiotic use. Moreover, we find that the selective pressure from antibiotics arises not so much from the rate of antibiotic use per se but from the frequency of treatment: frequent antibiotic therapy disproportionately impacts the fitness of sensitive strains. This same phenomenon explains why serotypes with longer durations of carriage tend to be more resistant. These dynamics may apply to other potentially pathogenic, microbial commensals and highlight how population structure, which is often omitted from models, can have a large impact.

1. Introduction

Despite the enormous burden of antibiotic resistance, the ways in which microbial populations maintain different patterns of resistance are not well understood. The malaria parasite *Plasmodium falciparum* has steadily accumulated resistance to all major drugs used against it [1], whereas *Streptococcus pneumoniae* gains and loses resistance over space and time, at least partly in response to local selection pressure [2,3]. These different dynamics are not yet predictable. It is intuitive that the spread of resistance depends on its genetic determinants and their fitness costs as well as the strength of selection from treatment. But it is also apparent that the spread of resistance changes with the strength of competition between resistant and sensitive strains within hosts, and between resistant and sensitive strains in the host population [4–6]. These interactions can be influenced by treatment practices and host behaviour. Effectively managing resistance in microbial populations requires understanding how these dynamics shape the competitive balance between resistant and sensitive strains.

The maintenance of antimicrobial resistance in *S. pneumoniae*, or pneumococcus, presents such a theoretical and practical challenge. Resistant strains seem to be maintained at stable frequencies in different populations, and yet simple models predict that coexistence should be rare, and either sensitive or resistant...
strains should fix [7]. The patterns suggest balancing selection, whereby coexistence is maintained by negative frequency dependence in the fitness of individual strains. The fraction of strains of *S. pneumoniae* that are non-susceptible to penicillin correlates strongly with local treatment rates among countries in Europe [8–10], provinces in Spain [11] and states in the USA [12,13]. Additionally, periods of high resistance correlate with periods of high local antibiotic usage [14]. These trends suggest that the frequency of non-susceptible strains is driven by treatment, which is consistent with studies showing that amoxicillin therapy disproportionately speeds the clearance of sensitive strains relative to non-susceptible strains [15]. Furthermore, the fact that resistance consistently declines during periods of low antibiotic usage suggests that resistance carries a fitness cost [14,16,17]. Sensitive strains tend to outcompete resistant strains in experimentally infected rats in the absence of antibiotics [18]. But the relative advantage of resistant strains in treated hosts, and of sensitive strains in untreated hosts, cannot by itself explain balancing selection for resistance in pneumococcus. Parsimonious transmission models show that resistant and sensitive strains can coexist in only a very narrow range where treatment pressure nearly equals the fitness cost of resistance [7]. Prescription rates of penicillins vary over threefold between countries in Europe [2,8], and the high rates of recombination of resistance elements in the pneumococcal genome suggest the fitness cost may vary between strains [19], a hypothesis also consistent with evidence from *in vitro* and animal models [18,20].

Different forms of competition could mediate selection for resistance and expand the potential for coexistence. In theoretical models, increasing the segregation of sensitive and resistant strains, such as by introducing explicit classes of treated and untreated hosts, slightly increases the range of outcomes in which both strains coexist [7]. Models with greater host population structure similarly predict more coexistence between resistant strains of *Staphylococcus aureus* that are adapted to community versus hospital transmission, although this structure generally does not enable the persistence of sensitive strains [21]. Allowing strains to compete more strongly with themselves than with one another also facilitates coexistence [7]. This kind of competition, mediated by specific immunity, is probably important for maintaining the diversity of pneumococcal serotypes [22]. Serotype variation may be one driver of coexistence, because the fitness benefit of resistance will be larger in serotypes of longer duration of carriage (and thus higher likelihood per carriage episode of antibiotic exposure) [23]. Indeed, prevalence of resistance is higher in more common serotypes, which have longer duration of carriage [24–26]. This mechanism can be generalized: modelling suggests that any phenotype that is simultaneously under balancing selection and associated with variation in duration of carriage will, by epidemiologically generated linkage, be associated with antibiotic resistance [23]. Candidate mechanisms in addition to serotype, such as phage content, have been identified [27], but it is still unclear how much variation in resistance this effect explains, especially within serotypes. Here, we focus instead on quantifying the impact of effects related to population structure.

To investigate the forces maintaining balancing selection on resistant and sensitive strains, we modelled the competitive dynamics of pneumococcus from the scale of individuals to communities. Resistant strains of pneumococcus are ubiquitous in human populations and arise repeatedly through recombination, and so we focus on the effects of treatment and host behaviour rather than the evolution of resistance *de novo*. Our results uncover new mechanisms contributing to balancing selection, including an unexpected role of treatment frequency, and explain differences in resistance rates between serotypes. Such insights suggest new areas of focus for managing pneumococcal resistance.

2. Results

We used an individual-based model to simulate the transmission dynamics of 25 pneumococcal serotypes in a host population whose age structure and vital rates were similar to those of the USA (Material and methods). Serotypes differed in their intrinsic fitnesses, with some serotypes having longer duration of carriage and greater within-host competitive ability, and thus higher prevalence, than others. Serotypes occurred as resistant and sensitive strains. Sensitive strains had longer durations of carriage in untreated hosts, and resistant strains had longer durations of carriage in treated hosts, but there was otherwise no difference between strains of a given serotype. Although more than 90 serotypes have been recorded, a smaller number are typically present in any given sample and we have selected 25 as a tractable number that is capable of capturing the observed diversity. Rates of clearance in treated hosts were fixed for both strains based on experimental values [15]. Treatment occurred independently of the host’s carriage status. To mitigate the effects of extinction due to drift in our artificially small population of 100,000 hosts, individuals were exposed to immigrant strains at approximately 0.1% of the rate at which they were exposed to strains from within the population. We examined the consequences of varying the fitness cost of resistance, defined as the per cent reduction in resistant strains’ duration relative to sensitive strains’ duration in untreated hosts, and the intensity of treatment, which was normalized to the USA rate of penicillin use. For the different models, we examined the range of fitness costs in which resistant and sensitive strains could coexist.

2.1. High antibiotic pressure alone can only select for resistance with modest cost

We started with a randomly mixing population in which there is no difference among hosts in the probability of receiving treatment. Hosts vary in their treatment statuses over time, creating a shifting pool of habitats favourable to the survival and transmission of either sensitive or resistant strains. In simpler models, including discrete treated and untreated host compartments has slightly promoted coexistence [7]. In our simulations, increasing the rate of antibiotic treatment predictably increased the fraction of resistant pneumococci. As expected, these increases shrank as the cost of resistance grew (figure 1). Over the range of treatment rates observed in Europe (corresponding to treatment levels of approx. 0.5–1.5 in our model), only very small fitness costs (less than or equal to 2%) were compatible with the coexistence of resistant and sensitive strains. This assumes a broad definition of coexistence, where resistant and sensitive strains each have frequencies between 2% and 98%. We further require for coexistence that the fraction of resistant strains increases at least 10% over this range of treatment rates.
(e.g. if 40% of strains are resistant at the 0.5 treatment level, at least 44% should be resistant at the 1.5 treatment level) to match the observed positive correlation between resistance and treatment [8–13]. Because epidemiological and experimental evidence suggests the cost may exceed 2% [14,18], we conclude that a randomly mixed, equally treated population cannot maintain stably coexisting penicillin-sensitive and -resistant pneumococci given mean penicillin prescription rates, despite the natural discretization of treated and untreated individual hosts.

Figure 1. Fraction of all pneumococci resistant by treatment level and fitness cost for the model with identical treatment rates by age and random mixing. Each coloured point shows the mean from one replicate simulation. Grey points show resistant fractions observed in European countries [2].

2.2. Structure from transmission and treatment promote coexistence

We next evaluated the impact of three additional forms of host population structure, all biologically supported, that might promote the coexistence of resistant and sensitive strains at greater fitness costs. People preferentially contact others in their age group, and so we simulated transmission dynamics with age-assortative mixing drawn directly from contact surveys [28]. Assortative mixing slightly increases the rate of coexistence when resistance carries a cost (figure 2).

Antibiotic prescription rates are higher in children than in adults, especially for respiratory tract infections. Keeping the total rate of antibiotic usage in the population constant, we varied the treatment rates by age group to match the distribution of penicillin prescriptions in US populations (see electronic supplementary material, figure S1). Without treatment, the model predicts higher carriage rates in children than in adults due to their differences in immunity [22]. Shifting treatments to young children predictably increases the fitness of resistant strains, increasing the overall fraction of resistant pneumococci in the population as a whole and also differences between the frequency of resistance in children and adults (figure 3). Carriage rates remain higher in children than in adults even though children receive proportionately more antibiotics (see electronic supplementary material, figure S2). Compared with both the null model and a model of age-assortative mixing alone, a model with age-specific treatments extended the range of fitness costs compatible with coexistence of resistant and sensitive strains (less than or equal to 4%; figure 2).

We next considered the impact of host population structure over a larger scale. In the simulated population of 100,000 hosts, the less common serotypes occasionally go extinct. By default, we assume that all immigrating strains have a 1% probability of being resistant, regardless of serotype. A more plausible assumption is that immigrating strains have the same probability of being resistant as a random strain of that serotype in a population with typical treatment rates (Material and methods). With ‘pseudo-spatial immigration’, the range of costs supporting coexistence is narrower than with age-specific treatments, although coexistence is more likely at a 2% cost (figure 2). Because immigrating strains are involved in approximately 0.1% of colonization events on average, immigration’s main effect on coexistence is to reseed strain populations that have gone extinct. Small changes in the probability of resistance among immigrants can nonetheless exert a slight effect on the probability of coexistence at different costs.

Combined, these forms of host population structure increase the range of fitness costs that allow coexistence and the frequency of coexistence for any cost (figure 2). Increased segregation of treated and untreated hosts through age-assortative mixing and age-specific treatment expands the range of tolerated costs. Pseudo-spatial immigration increases the frequency of coexistence across a wider range of costs. These results were robust to some uncertainty in the duration of carriage of resistant strains in treated hosts (see electronic supplementary material, figures S3 and S4).

2.3. Cotransmission has little effect on coexistence

Contacts with multiply colonized hosts create correlations in the risks of contact (and thus colonization) with different strains. Correlated transmission of groups of resistant or sensitive strains might promote their persistence in treated and untreated hosts. In the default model, by contrast, the risk of contacting a strain is identical for all hosts, or for all hosts of the same age with age-assortative mixing. We investigated the impact of correlated risks by simulating contacts between individual hosts, allowing the colonized host to ‘challenge’ the recipient host with each of the colonized host’s strains (Material and methods). Models with cotransmission modestly affected the frequency of coexistence (see electronic supplementary material, figure S4). Compared with the null model, models with cotransmission did not expand the range of costs supporting coexistence. However, when there was already host population structure, cotransmission could slightly increase the frequency of coexistence at the same range of costs if multiply colonized hosts were assumed to be equally infectious as singly colonized hosts (that is, if the probability of contacting each of the strains was divided by the total number of strains carried by the donor host). The minor impact of cotransmission is unsurprising because over 80% of hosts in these models are colonized with only one strain.

2.4. Common serotypes are more resistant

When resistant and sensitive strains coexist, resistance is more common in the most fit serotypes (figure 4), with the most fit ones sometimes showing fixation of resistance and the least fit showing fixation of sensitivity. This pattern changes little with age (see electronic supplementary material, figure S5). Serotypes in our model all have the same transmission rate, but they vary in two fitness components—their maximum expected durations of carriage and their ability to exclude other colonizing strains and thus in prevalence. The most fit serotype persists, on average, 220 days in naive, untreated hosts, whereas the rarest serotypes persist only 25 days. More
2.5. Treatment incidence more important than prevalence for promoting resistance

For the same reason that treatment disproportionately affects the fitness of serotypes with long duration, increasing the incidence of treatment episodes—but leaving total antibiotic use for each age unchanged—could promote resistance in the pneumococcal population as a whole: shorter intervals between treatment events increase the number of colonizations with sensitive strains that can be disrupted, and thereby increase selection for resistance across all serotypes. Instead of a 10-day treatment duration, we simulated 2.5- and 5-day treatment durations while quadrupling and doubling the expected number of treatment events per capita, respectively, and also a 20-day treatment duration while halving the treatment frequency. Total antibiotic use still varied by age as before (see electronic supplementary material, figure S1), but, for each age, the mean duration determined whether there were more short treatments or fewer long treatments. Increasing the frequency of treatment reduced the mean fitness of sensitive strains, allowing resistant strains to persist at high fitness costs, even when antibiotic use was low. It also increased the fraction of replicate simulations producing coexistence across all fitness costs (figure 5; electronic supplementary material, figure S4).

3. Discussion

It is not easy to justify balancing selection for resistant and sensitive strains of pneumococcus from first principles. Simple models predict that sensitive strains should generally exclude resistant strains when fitness costs of resistance are high, and resistant strains should exclude sensitive strains when the cost is low [7]. This explanation is unsatisfying given the observed range of frequencies at which both resistant and sensitive strains appear to coexist, and given the likely variation in resistance cost. We find that host population structure is critical to strain persistence, as it creates semi-protected habitats (niches) in which each strain has a fitness advantage. Our analysis highlights another aspect that is missing from some simple models. Strain fitnesses depend not only on the overall prevalence of antibiotic treatment in the population but also on the frequency of treatment episodes. This dependence helps to explain why common serotypes, which have long carriage durations, tend to be resistant. An aspect not considered in this model is that shorter courses of a 10-day treatment duration, we simulated 2.5- and 5-day treatment durations while quadrupling and doubling the expected number of treatment events per capita, respectively, and also a 20-day treatment duration while halving the treatment frequency. Total antibiotic use still varied by age as before (see electronic supplementary material, figure S1), but, for each age, the mean duration determined whether there were more short treatments or fewer long treatments. Increasing the frequency of treatment reduced the mean fitness of sensitive strains, allowing resistant strains to persist at high fitness costs, even when antibiotic use was low. It also increased the fraction of replicate simulations producing coexistence across all fitness costs (figure 5; electronic supplementary material, figure S4).

3. Discussion

It is not easy to justify balancing selection for resistant and sensitive strains of pneumococcus from first principles. Simple models predict that sensitive strains should generally exclude resistant strains when fitness costs of resistance are high, and resistant strains should exclude sensitive strains when the cost is low [7]. This explanation is unsatisfying given the observed range of frequencies at which both resistant and sensitive strains appear to coexist, and given the likely variation in resistance cost. We find that host population structure is critical to strain persistence, as it creates semi-protected habitats (niches) in which each strain has a fitness advantage. Our analysis highlights another aspect that is missing from some simple models. Strain fitnesses depend not only on the overall prevalence of antibiotic treatment in the population but also on the frequency of treatment episodes. This dependence helps to explain why common serotypes, which have long carriage durations, tend to be resistant. An aspect not considered in this model is that shorter courses of a 10-day treatment duration, we simulated 2.5- and 5-day treatment durations while quadrupling and doubling the expected number of treatment events per capita, respectively, and also a 20-day treatment duration while halving the treatment frequency. Total antibiotic use still varied by age as before (see electronic supplementary material, figure S1), but, for each age, the mean duration determined whether there were more short treatments or fewer long treatments. Increasing the frequency of treatment reduced the mean fitness of sensitive strains, allowing resistant strains to persist at high fitness costs, even when antibiotic use was low. It also increased the fraction of replicate simulations producing coexistence across all fitness costs (figure 5; electronic supplementary material, figure S4).
of antimicrobial use. Overall, our results suggest that managing resistance in pneumococcus should include consideration of treatment frequency, not just its prevalence, as well as the dynamics of transmission and treatment in core groups such as children.

Aggressive chemotherapy can prevent the de novo emergence of resistant strains if drug-sensitive strains are suppressed sufficiently rapidly [4,5], but when relatively fit resistant strains are already in circulation aggressive chemotherapy can remove drug-sensitive competitors and speed the selection and emergence of resistant strains. Selection in our model proceeds by the latter mechanism. By rapidly clearing sensitive strains, treatment reduces their fitness advantage in transmission, promoting the growth of resistant strains in the community. This is consistent with epidemiological studies, which show that the clearance of sensitive strains following treatment appears to have a stronger effect on individual risk of carriage than competitive release of resistant strains [31]. In other words, although treated hosts are at increased risk of colonization with resistant types, the fraction of resistant strains is mostly driven by changes in the abundance of sensitive strains. A corollary is that the prevalence of treatment should correlate negatively with the overall prevalence of pneumococcus in the community. This pattern appeared in the model (see electronic supplementary material, figure S7) but has not, to our knowledge, been investigated in natural settings.

Resistant fractions above 60% are rarely observed in large populations [2,8], suggesting that our models cannot easily explain observed patterns: our simulated pneumococci are consistently too strongly affected by a threefold variation in selective pressure (see electronic supplementary material, figures S3 and S8). There are several additional mechanisms, not included in our model, that may produce coexistence over

Figure 4. Resistance in young children (less than 5 years) by serotypes’ fitness ranks for different costs of resistance in the model with age-assortative mixing, age-specific treatment and pseudo-spatial immigration. Treatment is held at 1.0, equivalent to US rates. Each point shows the mean fraction of each serotype that is resistant from one replicate simulation. Means were obtained by averaging the last 50 years.
show may explain some of the observed coexistence; our approach focused on isolating individual model components to understand each component in isolation, and in combination. Combining the effect of population structure with other mechanisms, such as genetic variation in carriage duration beyond the capsular locus [23] and genetic linkage between resistance to different antibiotics, may fully explain coexistence, but will also likely result in complex models that are hard to understand or validate. A parsimonious approach will focus on combining validated and quantified effects with known impact on antibiotic resistance prevalence.

This work demonstrates the importance of seemingly fine details of model structure in developing expectations about competing strains. Previous models based on ordinary differential equations considered several elements of our model separately, but they omitted key features. We suspect that the most important are the discretization of treatment to individual hosts (similar to a treated compartment) and individual periods, as well as the fine population and age structure from age-assortative mixing and age-specific treatment rates. We show the importance of treatment frequency, not merely treatment rate. The simultaneous colonizations with multiple strains allowed by our model affect competition within hosts; simultaneous transmission of dual infections was previously shown to be important for coexistence [7]. Investigating the dynamical differences between seemingly parsimonious models, and identifying the rates that determine, for example, the strength of competition in different scales, will be important for rationally managing resistance.

4. Material and methods
The first section contains a general overview of the model, whose core has been described previously [22]. Later sections describe components in depth. The parameters are in table 1.

4.1. Overview
We used an individual-based model that tracked the carriage histories and ages of hosts over multiple generations [22]. Twenty-five serotypes of pneumococcus, each with two resistance phenotypes (resistant or sensitive), were simulated. All serotypes had the same transmission rate, but they differed in their durations of carriage and ability to exclude other serotypes when colonizing a host. These two traits were positively correlated, so there was a consistent ranking of serotypes’ fitnesses. Hosts gained serotype-specific and non-serotype-specific immunity when colonizing a host. These two traits were positively correlated, so there was a consistent ranking of serotypes’ fitnesses. Hosts gained serotype-specific and non-serotype-specific immunity, the former reducing susceptibility to future colonizations and the latter incrementally shortening the duration of carriage. Immunity thus depended on the total number of prior pneumococcal colonizations and their serotype composition, and it was independent of strains’ antibiotic sensitivities.

Antibiotic resistance could have a cost. For each serotype, resistant strains had a duration of carriage in untreated hosts that was some constant fraction $\xi$ of the sensitive strains’ durations (in untreated hosts) [18]. Values of $\xi < 1$ imply an intrinsic cost of resistance. Because the cost of resistance is uncertain, a range of values of $\xi$ was tested, and the cost was expressed as a percentage $(1 - \xi) \times 100\%$.

Both resistant and sensitive strains were negatively affected by antibiotic treatment, but sensitive strains were impacted more. Clearance rates in treated hosts have been estimated as 0.32 d$^{-1}$ for sensitive strains and 0.14 d$^{-1}$ for resistant strains [15]. We fixed the mean clearance rate of sensitive strains in treated hosts 0.25 d$^{-1}$ (a mean duration of 4 days) and assumed

Figure 5. Fraction of replicate simulations producing coexistence for different treatment durations, which correlate inversely with treatment frequency. Parameters are identical between replicates. The mean prevalence of treated hosts is held constant across models, so 5-day treatment durations are associated with twice the number of treatment events as the 10-day default, and 20-day treatment durations are associated with half as many. All models include age-assortative mixing, age-specific treatments, pseudo-spatial immigration and cotransmission with equally infectious hosts. Coexistence is defined as a resistant fraction between 2% and 98% and a 10% increase in this fraction as the treatment intensity varies from 50% to 150% of typical rates. Error bars show 95% of the distribution of coexistence means from bootstrapping 1000 times over 20 triplets of runs at different treatment rates.

a wider range of parameter values in reality than our model can capture. With age-assortative mixing and age-specific treatment, the spread of resistant strains is mostly limited by the fitness cost of resistance and the intensity of treatment. Relative to natural populations, our simulated hosts have little spatial and demographic structure. Households, schools and regional transmission all influence pneumococcal dynamics. Because these structures amplify the segregation of treated, high-transmitting hosts (especially children), they should facilitate constrained spread of resistant strains. This is especially likely if antibiotic treatment tends to follow infections by respiratory viruses, which have spatially localized epidemics. Spatio-temporal variation with migration might explain why resistance rarely fixes in the most common serotypes, as our model sometimes predicts. A separate potential mechanism promoting coexistence is genetic linkage between resistance to $\beta$-lactams and other antibiotics [32]. Use of other antibiotics would thus select inadvertently for resistance to $\beta$-lactams. This may be an important factor driving resistance: trimethoprin-sulfamethoxazole and cephalosporin use can be better predictors of penicillin non-susceptibility in individual hosts and communities than penicillin use [11,12,33]. The effective rate of antibiotic treatment may thus be higher than predicted from penicillin prescriptions alone (which constitute approximately one-third to two-thirds of total antibiotic prescriptions in European ambulatory settings [10]), promoting the spread of resistant strains and increasing the range of tolerable fitness costs. Finally, if resistant elements induce adaptive immunity, or if they are linked with structures that do, strong intrastrain competition could promote coexistence [7]. In the same vein, variation in the capsular locus and other loci that influence the duration of carriage, such as phage [27], and that are also under negative frequency-dependent selection is expected to promote coexistence through epidemiologically generated linkage [23]. Carriage duration determines the benefit of resistance, and selection for different durations indirectly selects for varying degrees of resistance.

The purpose of this work has been to explore the effect of known heterogeneities in population structure, which we
### Table 1. Parameters in the models.

<table>
<thead>
<tr>
<th>parameter</th>
<th>description</th>
<th>value</th>
<th>rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>transmission rate</td>
<td>$0.045 d^{-1}$</td>
<td>see text</td>
</tr>
<tr>
<td>$w$</td>
<td>per capita immigration rate per serotype</td>
<td>$1.25 \times 10^{-7}$ days</td>
<td>see text</td>
</tr>
<tr>
<td>$p_z$</td>
<td>probability immigrating strain of serotype $z$ is resistant</td>
<td>1%</td>
<td>see text</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>age-assortative mixing weights</td>
<td>—</td>
<td>[28]</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>susceptibility reduction from serotype-specific immunity</td>
<td>0.3</td>
<td>[22,34]</td>
</tr>
<tr>
<td>$\gamma(z)$</td>
<td>intrinsic duration of sensitive strain carriage in untreated host</td>
<td>$25 - 220$ days</td>
<td>[22,35,36]</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>minimum duration of carriage in untreated host</td>
<td>25 days</td>
<td>[22,36]</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>shape parameter for non-specific immunity</td>
<td>0.25</td>
<td>[22,35]</td>
</tr>
<tr>
<td>$\xi$</td>
<td>fitness cost of resistance</td>
<td>$[0.9, 1]$</td>
<td>varied</td>
</tr>
<tr>
<td>$q$</td>
<td>initial fraction of colonizing strains resistant</td>
<td>0.5</td>
<td>see text</td>
</tr>
<tr>
<td>$u$</td>
<td>initial fraction of hosts colonized with each serotype</td>
<td>0.02</td>
<td>see text</td>
</tr>
<tr>
<td>$\mu_{\text{max}}$</td>
<td>reduction in susceptibility from carrying fittest serotype</td>
<td>$0.25$</td>
<td>[22,37]</td>
</tr>
<tr>
<td>$\gamma_s$</td>
<td>sensitive strain clearance rate with treatment</td>
<td>$1/4 d^{-1}$</td>
<td>[15]</td>
</tr>
<tr>
<td>$\gamma_j$</td>
<td>ratio of resistant to sensitive strains’ durations with treatment</td>
<td>4</td>
<td>varied</td>
</tr>
<tr>
<td>$\sigma_{\text{max}}$</td>
<td>per capita treatment rate (unnormalized)</td>
<td>$0.354 yr^{-1}$</td>
<td>varied</td>
</tr>
<tr>
<td>$\sigma_{\text{mean}}$</td>
<td>mean treatment duration</td>
<td>10 days</td>
<td>[15]</td>
</tr>
<tr>
<td>$\sigma_{\text{std}}$</td>
<td>standard deviation of treatment duration</td>
<td>3 days</td>
<td>—</td>
</tr>
<tr>
<td>$\mu$</td>
<td>minimum interval between treatments</td>
<td>1.5 days</td>
<td>—</td>
</tr>
<tr>
<td>$N$</td>
<td>number of hosts</td>
<td>100 000</td>
<td>—</td>
</tr>
<tr>
<td>$\text{PMF of age at death}$</td>
<td>—</td>
<td></td>
<td>[22,38]</td>
</tr>
<tr>
<td>$\alpha_{\text{init}}$</td>
<td>initial fraction of hosts immune</td>
<td>0.5</td>
<td>see text</td>
</tr>
<tr>
<td>$\alpha_{\text{colonized}}$</td>
<td>initial fraction of hosts colonized with each serotype</td>
<td>0.02</td>
<td>see text</td>
</tr>
<tr>
<td>$\alpha_{\text{resistant}}$</td>
<td>initial fraction of colonizations resistant</td>
<td>0.5</td>
<td>see text</td>
</tr>
</tbody>
</table>

resistant strains lasted four times as long. This clearance rate of resistant strains in treated hosts is lower than empirical estimates, but if resistant strains were cleared much faster antibiotic use dramatically depressed carriage prevalence in the population. Such a trend has not been reported.

To avoid spurious results from stochastic extinctions in our finite host population, we allowed each serotype to have a very small rate of immigration $w$. The rate of immigration $w$ was set equivalent to one individual in the population being exposed every 80 days to another individual outside the population carrying each serotype–strain combination. By default, immigrating strains had a 1% probability of being resistant, $p_z = 0.01$, regardless of serotype $z$. With pseudo-spatial immigration, 10 populations with average antibiotic usage rates (treatment = 1.0) were simulated for 150 years with $p_z$ set individually for each serotype to the current fraction of colonizations resistant in the population, or $p_z = 0.01$ if no colonizations for that serotype were present. The fraction of resistant strains in each serotype was then averaged over the last 50 years and over all simulations to calculate the long-term fraction of resistance in each serotype. These serotype-specific values of $p_z$ were then used in simulations with other rates of antibiotic use.

### 4.2. Transmission

At discrete time steps, the force of colonization on a host for each serotype $z$ and strain $i$ (i.e. a resistant or sensitive strain) in the simplest model was computed following:

$$
\Lambda(z, i) = q(z, \theta, C) \left( \beta \frac{i}{N} + wp_z \right). 
$$

The term $q(z, \theta, C)$ is the host’s susceptibility, or probability of acquiring serotype $z$, contingent on the vectors of past colonizations $\theta$ and current carriage $C$. Vectors $\theta$ and $C$ are both indexed by serotype. Entries of $\theta$ count the number of times the host has cleared each serotype, and $C$ counts the number of current colonizations with each serotype. The contact rate, $\beta$, is shared by all serotypes. The effective fraction of hosts colonized with serotype $z$ and strain $i$, $I_zi/N$, equals the number of colonizations with $(z, i)$ in the population divided by the population size $N$. This representation means that hosts who are colonized with a single strain of serotype $z$ are counted once, and hosts harbouring more than one strain of serotype $z$ are counted multiple times; this formulation avoids biasing the model towards coexistence [39].

In the model that includes age-assortative mixing, the force of colonization with serotype $z$, strain $i$, on a host of age $a$ is computed by

$$
\Lambda(a, z, i) = q(z, \theta, C) \left( \sum_{j=1}^{a} \alpha_{a} \beta \frac{I_{ij}}{N_{j}} + wp_z \right).
$$

Here, $\alpha_{a}$ weights contacts by individuals of age $j$ to individuals of age $a$. The matrix $\alpha$ of those weights is obtained from normalizing data from physical contact surveys from the UK and has single-year resolution [22,28].

### 4.3. Cotransmission

Two alternative modes of transmission test the impact of correlations in colonization risk with different strains. In each time step, the expected number of potential transmission events in the population is calculated as $\beta I$, where $I$ gives the total number of people with any pneumococcal colonization. For each contact, a random colonized host and random other host are selected. Under the ‘equally infectious hosts’ assumption, the receiving host is challenged with each colonization in the colonized
(donor) host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-

(\text{donor})$ host with probability $1/n$, where $n$ is the number of colo-


