**Within-host dynamics shape antibiotic resistance in commensal bacteria**

Nicholas G. Davies[1,2*], Stefan Flasche[1,2], Mark Jit[1,2,3], Katherine E. Atkins[1,2,4]

[1] Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.
[2] Department for Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.
[3] Modelling and Economics Unit, Public Health England, London SE1 8UG, UK.
[4] Centre for Global Health, Usher Institute of Population Health Sciences and Informatics, Edinburgh Medical School, The University of Edinburgh, Edinburgh EH8 9AG, UK.
* To whom correspondence should be addressed. E-mail: Nicholas.Davies@lshtm.ac.uk

**The spread of antibiotic resistance, a major threat to human health, is poorly understood. Simple population-level models of disease transmission predict that above a certain rate of antibiotic consumption in a population, resistant bacteria should completely eliminate non-resistant strains, while below this threshold they should be unable to persist at all. This prediction stands at odds with empirical evidence showing that resistant and non-resistant strains coexist stably over a wide range of antibiotic consumption rates. Not knowing what drives this long-term coexistence is a barrier to developing evidence-based strategies for managing the spread of resistance. Here, we argue that competition between resistant and sensitive pathogens within individual hosts gives resistant pathogens a relative fitness benefit when they are rare, promoting coexistence between strains at the population level. To test this hypothesis, we embed mechanistically-explicit within-host dynamics in a structurally-neutral disease transmission model. Doing so allows us to reproduce patterns of resistance observed in the opportunistic pathogens *Escherichia coli* and *Streptococcus pneumoniae* across European countries, and to identify factors that may shape resistance evolution in bacteria by modulating the intensity and outcomes of within-host competition.**

Antibiotic-resistant infections tend to be more common in populations that consume more antibiotics[1–3]. The explanation seems obvious: greater antibiotic use selects for more resistance. But capturing this pattern in an explicit model of disease transmission has been notoriously difficult[4]. The problem is that empirical observation suggests a gently rising, roughly linear relationship between consumption and resistance, with both resistant and sensitive (*i.e.*, non-resistant) strains coexisting over a 4- to 20- fold range of antibiotic treatment rates[1–3] (Fig. 1a). In contrast, simple models of disease transmission predict competitive exclusion[5]—that is, they predict that resistant strains will either disappear completely or spread to fixation, depending upon the rate of antibiotic consumption in a population (Fig. 1b–d). Although potential explanations for this discord between theory and observation have been proposed[4,6–8], a generalisable, biologically-explicit mechanism that accounts for widespread coexistence has yet to be identified. In short, despite the global public health threat of antibiotic resistance[9,10], we do not fully understand how resistance spreads in human populations.

We propose that within-host competition shapes resistance evolution and can promote widespread coexistence in commensal bacteria (*i.e.*, species that are normally part of the host microbiota, but which occasionally cause disease when they invade sterile sites). Mathematical models of resistant disease transmission routinely overlook within-host interactions between different bacterial strains, but commensal bacteria regularly cohabit with genetically- and phenotypically-distinct strains of the same[11–15] or different[16–18] species. Laboratory experiments have shown that resistant and sensitive microbes inhibit each other's growth when co-colonising the same host[19–22], suggesting that these distinct strains engage in exploitative competition[23] for host resources. Meanwhile, theory developed for malarial parasites[24] has proposed that within-host competition between co-colonising resistant and sensitive strains may interact with antimicrobial treatment to generate frequency-dependent selection[25,26] for resistance at the population level, promoting coexistence. We develop this theory, arguing that

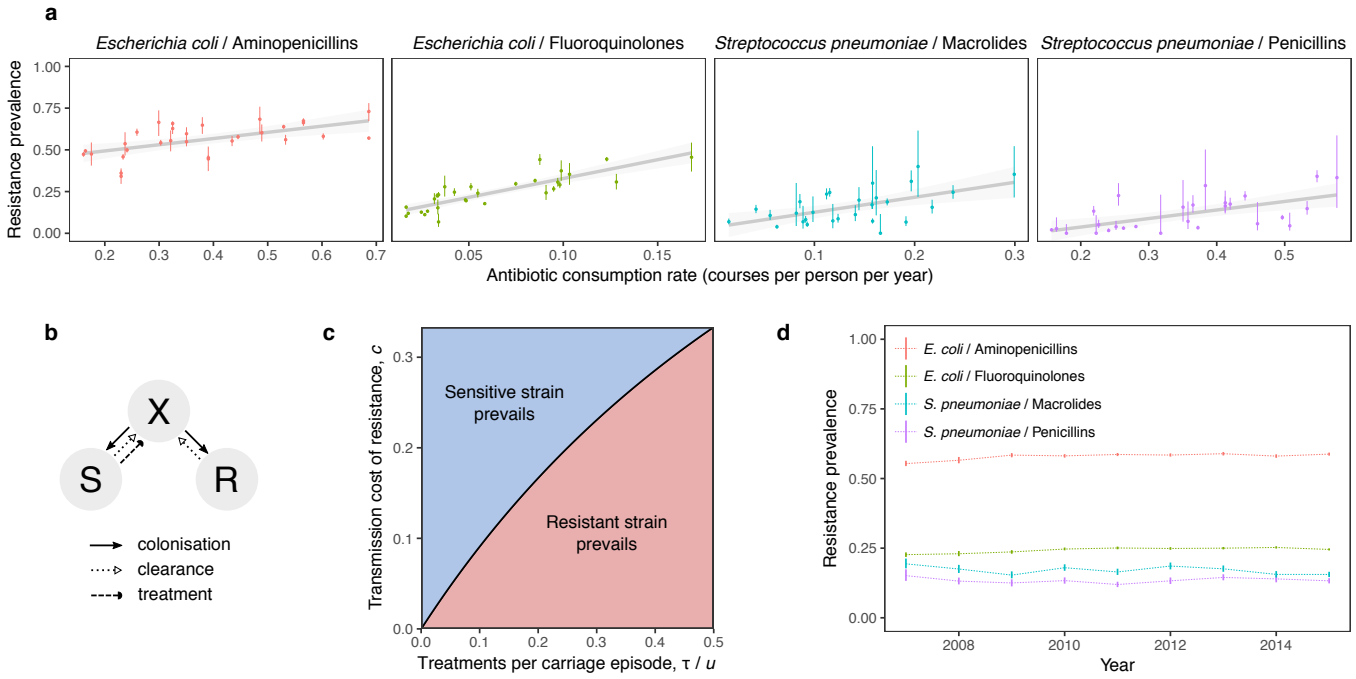



**Fig. 1 | The problem of coexistence. (a)** Resistant and sensitive strains of *E. coli* and *S. pneumoniae* coexist, and resistance increases moderately with antibiotic consumption[1–3]. The proportion of invasive isolates testing positive for drug resistance (with 95% confidence intervals) is plotted against community-level antibiotic consumption for 30 European countries (linear regressions with 95% confidence intervals overlaid). In contrast with observed coexistence, a simple model of resistant disease transmission **(b)** predicts competitive exclusion **(c)**. The model is defined by the system of ordinary differential equations $\frac{dS}{dt} = \beta SX - (u+\tau)S, \frac{dR}{dt} = \beta(1-c)RX - uR, X = 1 - S - R$, with $X$ non-carriers, $S$ sensitive-strain carriers, and $R$ resistant-strain carriers. Here, $\beta$ is the transmission rate (solid arrows), $u$ is the natural clearance rate (dotted arrows), $\tau$ is the antibiotic treatment rate (dashed arrow), and $c$ is the cost of resistance. When $\beta > u + \tau$ and $\beta(1-c) > u$, either strain can persist in isolation, but only one strain persists when both are present, with the sensitive strain prevailing when $\tau/u > c/(1-c)$. **(d)** The average resistance prevalence in Europe has hardly changed in recent years, suggesting that observed coexistence is stable rather than a transient state on the way to competitive exclusion.

population-level coexistence can be promoted by any phenotypic diversity that mediates competition between co-colonising strains. Accordingly, we expect co-colonisation to promote coexistence not only between resistant and sensitive bacteria, but also among other diverse microbes exploiting the same host niche, such as pneumococcal serotypes[27].

We develop a "mixed-carriage" model that mechanistically captures within-host competition in an explicit model of bacterial transmission. This stochastic individual-based model—which can be approximated using deterministic ordinary differential equations (ODEs) for analytical simplicity—observes the key requirement of structural neutrality[28], *i.e.*, it avoids systemic biases that non-mechanistically promote (or inhibit) coexistence. When fit to data across 30 European countries, the model provides a parsimonious and generalizable explanation for empirical patterns of resistance across four pathogen-drug combinations. We also show how within-host competition can help to explain observed patterns of resistance[7] and antigenic diversity[27] among competing serotypes of the commensal bacterium *Streptococcus pneumoniae*.

## Results

**Co-colonisation creates frequency-dependent selection for resistance.** Frequency-dependent selection[25,26] is known to promote diversity among competitors in animals[25,29], plants[30], and microbes[31]. In the classic scenario, a rare mutant invades a population by exploiting some weakness of wild-type individuals, but gradually becomes a victim of its own success by displacing the competitors it relies upon to exploit. Stable coexistence between types can result if mutants tend to increase in frequency when they are rare (because there are ample wild-type individuals to exploit) but decrease in frequency when they are common (because there are too few wild-type individuals to exploit). Extending a hypothesis suggested by Hastings for malarial parasites[24], we suggest that frequency-dependent selection for resistant bacteria is created by within-host competition among co-colonising strains.

The mechanism works as follows. Suppose that a small group of resistant cells could colonise one of two hosts. One host already carries sensitive bacteria, while the other carries resistant bacteria. All else equal, the resistant cells would benefit more by colonising the sensitive-cell carrier, because if that host were to subsequently take antibiotics—eliminating the resident sensitive cells—the newly-arrived resistant cells could multiply to fully exploit the host niche, increasing their potential to be transmitted to new hosts. Indeed, *in vivo* studies have shown that in co-colonised hosts harboring both sensitive and resistant cells, the resistant pathogens increase in abundance when their sensitive competitors are killed by antibiotic treatment[19,20,32,33]—that is, treatment results in competitive release[20] for the resistant cells. On the other hand, co-colonising the resistant-cell carrier offers no such benefit to resistant cells, because later antibiotic use gives no advantage to the invading bacteria over the resident bacteria. This disparity creates frequency-dependent selection for resistance (Fig. 2a) because—on average—a resistant cell is more likely to find itself co-colonising a sensitive-strain carrier when resistance is rare.

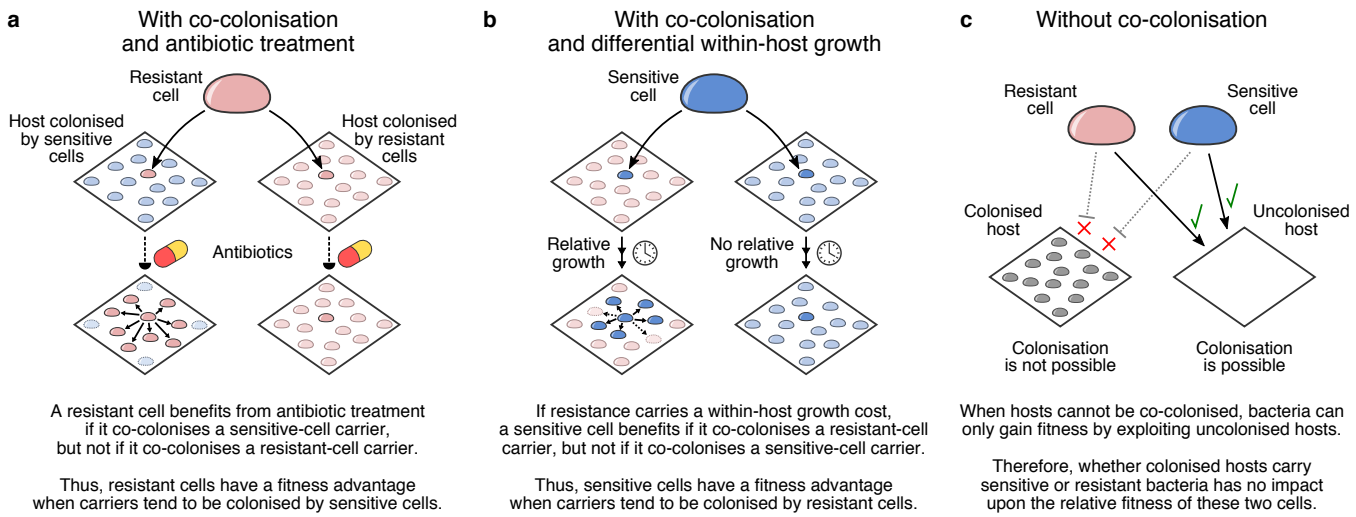



**Fig. 2 | Co-colonisation creates frequency dependent selection for resistance.** With co-colonisation, **(a)** antibiotic treatment causes resistant cells to have higher fitness when sensitive-strain hosts are more common and **(b)** differential within-host growth causes sensitive cells to have higher fitness when resistant-strain hosts are more common. Either mechanism can promote coexistence between resistant and sensitive strains. **(c)** Without co-colonisation, the relative frequency of sensitive-strain and resistant-strain carriers has no differential impact upon the fitness of resistant versus sensitive cells, so there is no frequency-dependent selection acting on resistance phenotypes.

Although originally phrased in terms of competition between malarial parasites mediated by antibiotic treatment and resistance[24], this mechanism has broader applicability. First, other forms of within-host competition—not just treatment-mediated competitive release—can promote coexistence. For example, *in vitro*[21] and *in vivo*[22,32] studies have shown that, in the absence of antibiotics, sensitive cells often exhibit greater within-host growth relative to resistant cells—consistent with resistance carrying a fitness cost[34,35] manifesting as a reduced growth rate. Sensitive cells would then benefit more from co-colonising a resistant-strain carrier than a sensitive-strain carrier (Fig. 2b). This relative advantage may also promote frequency-dependent selection acting on resistance phenotypes, because a sensitive cell is more likely to co-colonise a resistant-strain carrier when resistance is common. Second, there is no requirement that competing strains are closely related—only that they competitively suppress each other when colonising the same niche—although we focus here on competition between strains of the same species.

**Implicit versus explicit models of within-host dynamics.** Models that do not account for within-host competition will fail to capture this source of frequency-dependent selection for resistance (Fig. 2c). Nonetheless, existing models that do incorporate co-colonisation have not convincingly reproduced empirically-observed coexistence[4,6]. We suggest that these models have fallen short not because within-host competition is a poor driver of coexistence, but because they feature unrealistic assumptions concerning within-host dynamics. To illustrate this point, we compare two models of resistant disease transmission: an existing model[4] which we refer to as the "knockout model", and a new "mixed-carriage model". These models share the same population-level dynamics, but differ in how they capture within-host dynamics, resulting in a substantial disparity in population-level patterns of resistance.

The shared assumptions of both models are as follows. There are two co-circulating bacterial strains, one resistant and one sensitive. Hosts mix randomly, with each colonised host infecting other hosts at rate $\beta$, transmitting a "germ" to a randomly-selected host. A germ contains cells of one strain, chosen randomly in proportion to the number of cells of each strain carried by the transmitting host; all colonised hosts, including those carrying multiple strains, are assumed to be equally infectious. Resistant germs fail to transmit with probability $c$, where $c$ is the transmission cost of resistance[34,35]; additionally, transmission only succeeds with probability $k$ if the recipient is already a carrier, where $k$ is the efficiency of co-colonisation relative to primary colonisation. Finally, each host is naturally cleared of all strains at rate $u$, and cleared of sensitive cells by antibiotic treatment at an additional rate $\tau$.

Starting from this common framework, the two models make divergent assumptions about within-host dynamics. First, the existing "knockout" model[4,28] assumes that hosts can be treated as though they contain two subcompartments of equal size (Fig. 3a). When a germ is transmitted to an uncolonised host, the invading strain fills the entire host niche, occupying both subcompartments. If instead, germs are successfully transmitted to an already-colonised host, the invading strain "knocks out" and replaces the contents of one of the two subcompartments at random. These assumptions allow the knockout model to be implemented using only four host states—namely, X hosts are

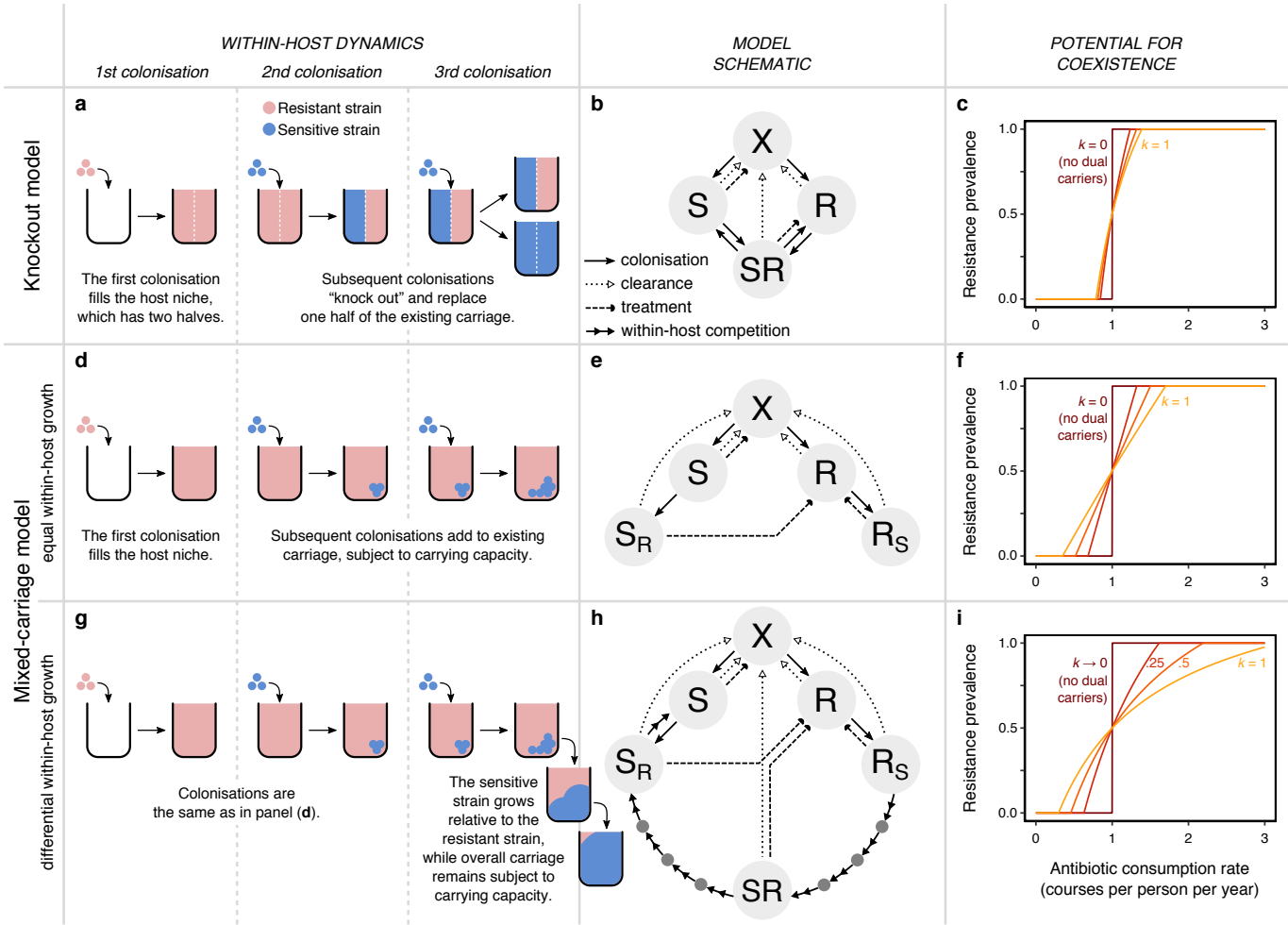



**Fig. 3 | Two models of within-host dynamics. (a)** In the knockout model[28], hosts contain two subcompartments. A resident strain must be "knocked out" from its subcompartment for a new strain to invade. **(b)** The knockout model requires four host states, adding the "SR" dual carriage state to the model of Fig. 1b. **(c)** We plot the equilibrium resistance prevalence (the probability that a randomly-selected pathogen from a randomly-selected host is resistant) as a function of the treatment rate $\tau$ and the relative efficiency of co-colonisation $k$ (with $k$ = 0, 0.25, 0.5, 1.0 shown, from dark to light). Coexistence increases with $k$ but remains limited. Setting $k$ = 0 recovers the single-strain model of Fig. 1b and competitive exclusion. **(d)** The mixed-carriage model explicitly tracks within-host strain frequencies and treats cells of either strain equally, relaxing the assumption of host subcompartments that contain only one strain at a time. When new cells enter the host, they mix freely with existing strains. **(e)** The mixed-carriage model can be approximated using five host states, where $S_R$ and $R_S$ represent hosts colonised primarily by one strain, with a small complement of the other. **(f)** Explicitly tracking within-host dynamics promotes coexistence. **(g)** We extend the mixed-carriage model to incorporate differential within-host growth of strains, adding **(h)** "intermediate" host states representing different relative frequencies of the two strains. Treatment and clearance events for intermediate states (dark grey circles) are omitted for clarity. **(i)** Within-host growth further promotes coexistence. In panels c, f, and i, $\beta$ = 5 mo$^{-1}$ and $u$ = 1 mo$^{-1}$, while specific values of $c \approx 0.07–0.12$ (panels c, f) and $w_s \approx 14–34$ (panel i) are chosen such that resistance prevalence passes through 0.5 when $\tau$ = 1 y$^{-1}$ (Supplementary Note 1).

uncolonised, S hosts carry the sensitive strain only, R hosts carry the resistant strain only, and SR hosts carry both strains, one in each subcompartment (Fig. 3b). In the Methods, we describe how these model dynamics may be analysed either using stochastic individual-based methods or by integrating systems of ordinary differential equations (ODEs).

As shown by Lipsitch *et al.*[28], the knockout model is the simplest mathematical model that allows co-colonisation without exhibiting systemic biases that artificially promote coexistence (*i.e.*, it is structurally neutral[28]). Nonetheless, a mechanistic interpretation of a host's two equally-sized subcompartments, as posited by this model, is challenging. For example, they could represent two physically-distinct but ecologically-equivalent niches, but the identity of these two niches would be unclear, and it is known that bacteria of different strains can readily occupy the same host niche[11–14]. Alternatively, the two subcompartments may be a way of representing a single host niche—*e.g.* the nasopharynx or the gut—but it is unclear why a group of invading cells should replace either all resistant cells or all sensitive cells from an SR carrier rather than replacing cells from either strain at random. In addition to these conceptual difficulties, the knockout model predicts coexistence only across a narrow range of treatment rates that does not reflect the wide range over which coexistence is observed empirically (Fig. 3c).

To overcome these issues, we propose a new "mixed-carriage" model that explicitly tracks within-host strain frequencies without splitting the host niche into two subcompartments. As in the knockout model, when a host is newly colonised, the invading strain is assumed to immediately occupy the entire host niche, reaching the host's carrying capacity (Fig. 3d). But when new cells enter, they are simply added to the cells that are already being carried. Carrying capacity is then immediately reimposed by eliminating excess cells at random, rather than by eliminating all cells from a given subcompartment containing only one strain. That is, following co-colonisation, the host niche contains a fraction $\frac{1}{1+\iota}$ of the "old" cells—an unbiased sample of the host's carriage prior to co-colonisation—and a fraction $\frac{\iota}{1+\iota}$ of the "new" cells, where $\iota$ is the "germ size", the relative size of an invading group of cells compared to the host's carrying capacity. Because this model allows hosts to carry an arbitrary mix of cells of different strains, it requires keeping track of a large number of host states, which our stochastic individual-based implementation achieves. However, under the simplifying assumption that germ sizes are small ($\iota \ll 1$), the model is well approximated using a system of ODEs with only five host states (Fig. 3e), for a similar mathematical tractability to the knockout model (see Supplementary Note 1 for details). Strikingly, the mixed-carriage model supports much more coexistence than the knockout model, suggesting that a more explicit model of within-host dynamics may more readily explain observed patterns of resistance (Fig. 3f).

Because it specifically tracks within-host strain frequencies, the mixed-carriage model can serve as a starting point for more complex models. To illustrate this, we add differential within-host growth to the model, such that sensitive cells gradually grow in frequency relative to resistant cells sharing the same host (Fig. 3g). Accordingly, we assume that the sensitive strain grows exponentially relative to the resistant strain at

rate $w_s$—eliminating the resistant strain completely if its relative within-host frequency drops below a critical threshold $f_{min}$—while overall carriage remains fixed at the host's carrying capacity. Again, this differential growth requires tracking a large number of host states, which can either be accounted for directly with an individual-based model implementation or be approximated using a finite number of mixed-carriage states in a system of ODEs, with the number of states depending upon the desired degree of concordance with the idealised dynamics of within-host growth (Fig. 3h; Supplementary Note 1). Differential within-host growth tends to gradually eliminate resistant cells from co-colonised carriers, partially reducing the frequency-dependent benefit associated with resistant cells co-colonising sensitive-strain carriers. However, it also introduces an additional frequency-dependent advantage for sensitive cells co-colonising resistant-strain carriers, which, overall, can further expand the potential for coexistence (Fig. 3i).

In each model, the potential for coexistence depends upon the prevalence of co-colonisation, which is partly governed by the parameter $k$: while setting $k = 0$ eliminates co-colonisation and recovers competitive exclusion, allowing co-colonisation ($k > 0$) promotes coexistence. In Supplementary Note 2, we identify the key processes that inhibit coexistence in the knockout model and promote coexistence in the mixed-carriage model, showing how the extent of coexistence depends crucially upon the prevalence of hosts carrying both sensitive and resistant strains.

**Structural neutrality of the knockout and mixed-carriage models.** A structurally-neutral model is one in which, when the biological differences between two strains are stripped away, pathogens of either strain are not treated differently from one another[28]. The aim of structural neutrality is to ensure that the predicted outcome of competition between strains—whether it is coexistence or competitive exclusion—is attributable to identifiable, biological differences between the strains, rather than to hidden assumptions embedded in the model structure. The knockout model meets the mathematical criteria for structural neutrality proposed by Lipsitch *et al.*[28], but we argue that it violates the spirit of neutrality nonetheless. Specifically, the knockout model assumes that when a host carrying pathogens of two different strains is invaded by a new strain, the invading strain completely replaces one of the two resident strains while leaving the other untouched—even if the two resident strains differ only by a neutral, biologically-meaningless label. This property artificially depletes within-host strain diversity, inhibiting coexistence by reducing the scope for within-host competition. By contrast, the mixed-carriage model avoids this artificial loss of diversity, while adhering to both the spirit and the letter of structural neutrality. In Supplementary Note 3, we demonstrate the structural neutrality of the mixed-carriage model, and discuss how a model's adherence to within-host neutrality depends upon the interpretation of within-host states.

**Explicitly capturing within-host dynamics reproduces widespread coexistence.** We used Bayesian inference via Markov chain Monte Carlo (MCMC) to fit both the knockout and mixed-carriage models to consumption and resistance data reported by 30 European countries across two common drug classes for the commensal pathogens *E. coli* and *S. pneumoniae*[2,3]. We assumed that countries differ only in antibiotic consumption, while other epidemiological parameters are shared across countries and

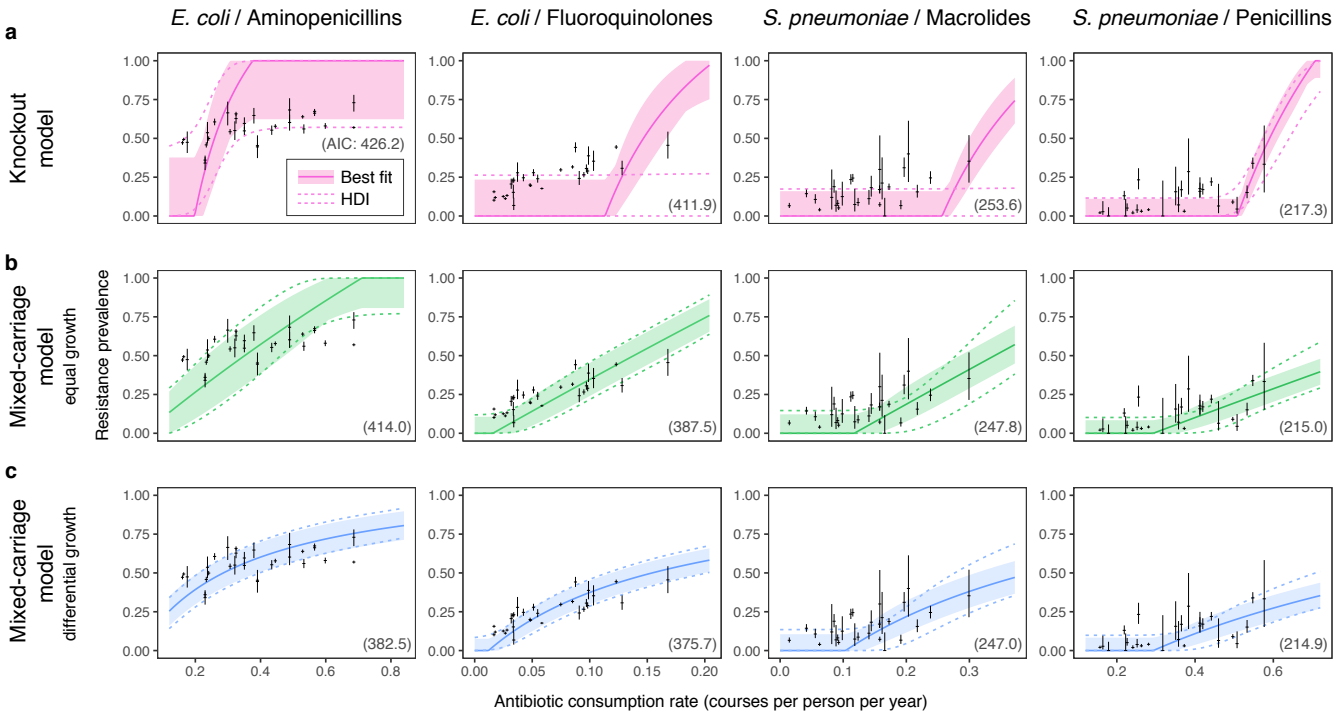



**Fig. 4 | Within-host dynamics explain patterns of resistance in commensal bacteria.** The knockout model **(a)** does not capture widespread coexistence, while the mixed-carriage model with **(b)** or without **(c)** differential within-host growth does. Solid lines and ribbons show the single best-fit run for each model (solid lines) and the 67% highest density interval (HDI) incorporating between-country random effects (shaded ribbon). Regions bounded by dashed lines show the 67% HDI across the estimated posterior, again incorporating between-country random effects. The Akaike Information Criterion associated with each model fit is given in parentheses on each panel; note that AICs are not strictly comparable across pathogen-drug data sets (columns) as the number of countries and sample sizes differ.

are constrained to be consistent with empirically-observed ranges for carriage prevalence and average duration of carriage. Due to the limited range of coexistence predicted by the knockout model, we find that it cannot capture observed patterns of resistance[4,6] (Fig. 4a). However, the empirical data are better captured by the mixed-carriage model (Fig. 4b), particularly when differential within-host growth is introduced (Fig. 4c). Using the Akaike Information Criterion to select the most parsimonious model, we find that the mixed-carriage model with differential within-host growth has the most statistical support across all bacteria-drug combinations (Fig. 3). Frequent co-colonisation by sensitive and resistant cells—irrespective of the overall prevalence of the species of interest—is needed to maintain widespread coexistence via within-host competition (Supplementary Note 4).

**Patterns of coexistence among pneumococcal serotypes.** So far, we have focused on a simplified scenario in which bacterial diversity is limited to sensitive versus resistant strains, but the mixed-carriage model can be extended in this respect. The nasopharyngeal coloniser *S. pneumoniae* exhibits extensive diversity in the expression of capsular proteins exposed to the host immune system, subdividing the species into nearly 100 distinct "serotypes" that—like resistant versus sensitive strains—are known to stably coexist in host populations[27,36]. Understanding both the coexistence of these serotypes and the evolution of resistance within each is vital for building a comprehensive picture of resistance evolution in pneumococci. We thus extended the two-strain mixed-carriage model (Supplementary Note 5) by parameterising it with the serotype-specific duration of carriage for 30 of the most common *S. pneumoniae* serotypes[7], assuming a 10% transmission cost and a 20% growth cost of resistance, and introduced serotype- specific adaptive immunity to the model (*i.e.* host immunity to colonisation by previously-cleared serotypes). The extended model captures much of the observed serotype diversity and patterns of resistance among serotypes (Fig. 5).

**General predictions of the mixed-carriage model.** Our extended serotype model illustrates that within-host competition can promote pathogen diversity more broadly than for resistance-associated phenotypes *per se*. For example, consider a host carrying cells of two different serotypes. If one serotype is cleared by the host immune system, the other serotype may benefit from competitive release. So long as clearance of one serotype does not result in clearance of all serotypes within a host, clearance will tend to promote rare serotypes, since the hosts they co-colonise are more likely to be carrying a different serotype, and hence they are more likely than common serotypes to be the beneficiaries of competitive release mediated by natural clearance. This effect can promote serotype diversity (Fig. 6a) even in the absence of any host acquired immunity[27,36].

We conclude by considering the impact of carriage duration, transmission rate and growth rate upon resistance evolution. In agreement with previous theoretical work[7], we find that a longer duration of carriage promotes greater resistance when resistance carries a transmission cost (Fig. 6b). However, this association can be reversed when resistance instead carries a within-host growth rate cost (Fig. 6c), because a longer duration of carriage affords sensitive cells a greater opportunity to outcompete resistant cells within hosts. Accordingly, the overall relationship between duration of carriage

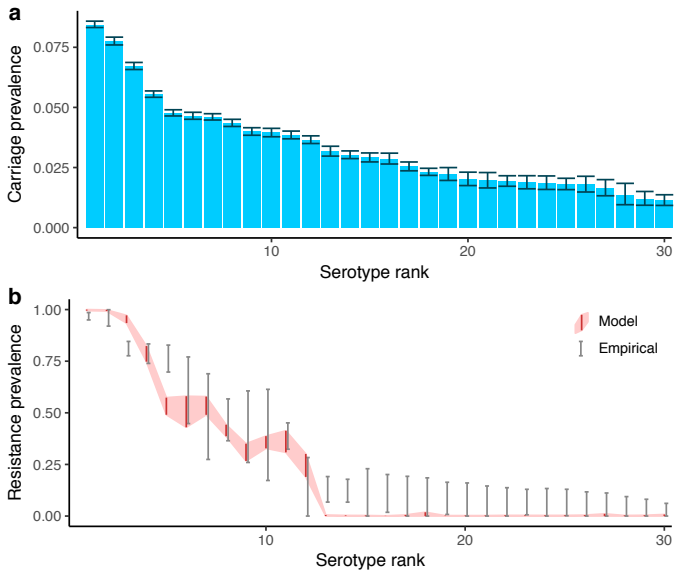



**Fig. 5 | Resistance in coexisting pneumococcal serotypes.** We use the mixed-carriage model to simulate 30 co-circulating pneumococcal serotypes, using a previously-published data set[7,50] to assign measured durations of carriage to each serotype, while incorporating a simple model of host adaptive immunity. We assume that serotypes with a longer duration of carriage also have a within-host growth rate advantage[27], and that resistance carries a 10% transmission cost and a 20% within-host growth cost. We recover extensive diversity in **(a)** pneumococcal carriage (error bars show 95% interquantile range for the prevalence of each serotype among carriers in the final 100 years of the 400-year simulation) and **(b)** resistance prevalence (grey error bars show 95% confidence intervals for empirical resistance prevalence; red ribbon shows 95% interquantile range for model resistance prevalence). Note that model serotypes are ranked from high to low duration of carriage (a, b) while empirical serotypes are ranked from high to low resistance prevalence (b), to facilitate comparing general trends of within-serotype coexistence. Results from one model run are shown.

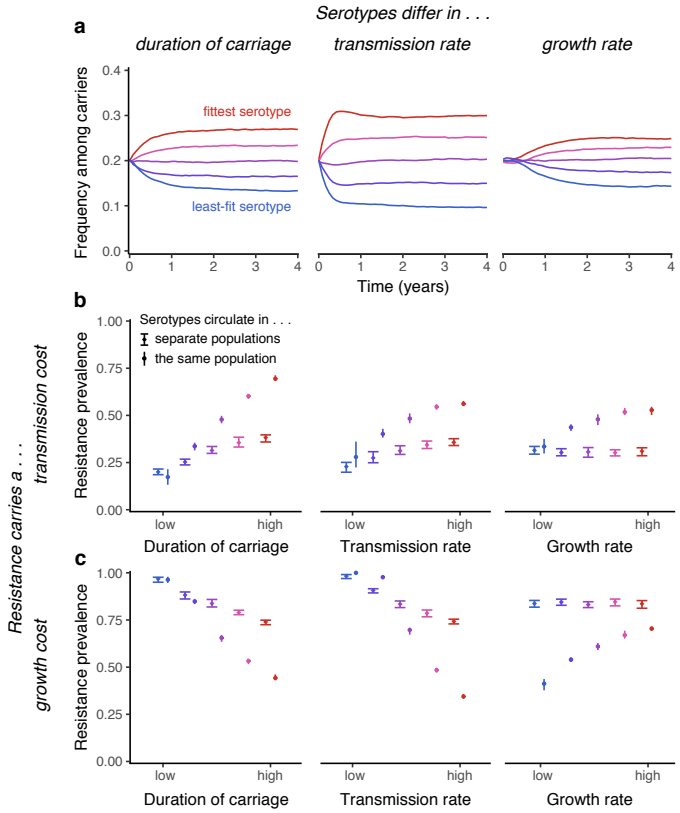



**Fig. 6 | General effects of within-host competition.** Serotype-specific clearance promotes coexistence between serotypes, and intrinsic fitness differences between serotypes are correlated with resistance prevalence within serotypes. Serotypes are assumed to differ in duration of carriage ($u$ = 1.04, 1.02, 1, 0.98, 0.96), transmission rate ($\beta$ = 1.84, 1.92, 2, 2.08, 2.16), or within-host growth rates ($w$ = 1, 2, 4, 8, 16). In each plot, the fittest serotype is shown in red. **(a)** In a model with five serotypes (all antibiotic-sensitive) differing in various measures of intrinsic fitness, serotype-specific clearance maintains coexistence between serotypes in the absence of any acquired immune response. **(b)** When resistance carries a 10% transmission-rate cost, fitter serotypes are more strongly selected for resistance. We contrast trends in resistance when serotypes circulate in separate populations or together in the same population; circulating together tends to magnify differences in resistance between serotypes. The mean and 95% interquantile range for the last 50 years of each 100-year simulation is shown. **(c)** When resistance carries a growth-rate cost (with sensitive strains growing at 10 times the rate of resistant strains), fitter serotypes are less strongly selected for resistance, except when serotypes differ in growth rate, where the trend is reversed. While serotypes circulating in the same population tends to increase average resistance prevalence when resistance carries a transmission cost, it tends to decrease resistance when resistance carries a growth cost. For each plot, results from a single model run are shown.

and resistance likely depends upon the balance of these two costs of resistance for a given species. Our model also predicts that a higher transmission rate promotes co-colonisation. In co-colonised hosts, sensitive strains may be eliminated by treatment, while resistant strains may be eliminated by faster-growing sensitive strains. The relative importance of these two forms of competition determines whether increased transmission promotes or inhibits resistance (Fig. 6b & c). This mechanism may elucidate an observed positive relationship between resistance prevalence and population density[37]. Finally, we find that resistance is promoted in serotypes with greater within-host growth, as they are less likely to be excluded by other serotypes before antibiotic treatment results in their competitive release. Each of these three trends appears stronger when serotypes circulate in the same population than in different populations (Fig. 6b&c). Why various species exhibit different levels of resistance when faced with similar rates of antibiotic treatment is an outstanding problem in resistance evolution, which further analysis may help to resolve.

## Discussion

Our model provides two advances over previous work: it harmonises pathogen dynamics by mechanistically capturing both between-host and within-host processes, and it better captures empirical patterns of antibiotic resistance. We argue that frequency-dependent selection drives these patterns of resistance, and that explicitly tracking within-host dynamics helps to reproduce them.

In order for within-host competition to maintain substantial coexistence, a high proportion of hosts must be colonised by both resistant and sensitive bacteria. Co-colonised strains must also compete for transmission; models with co-colonisation but no competitive release do not produce extensive coexistence[6]. Empirical estimates suggest that dual carriage may be widespread. A study of *Staphylococcus aureus* carriage in children found 21% of carriers were colonised by both resistant and sensitive *S. aureus* strains[14]. Relatively few studies have measured simultaneous carriage of both sensitive and resistant strains of the same species, but carriage of multiple strains more generally appears to be common: genotyping studies have found up to 48% multiple carriage of genetically-distinct *S. pneumoniae* strains[11,12] and up to 86% multiple carriage of *E. coli* strains[13,15]. Although we have focused on competition between conspecific strains, competition between different species could also promote coexistence, reducing the need for widespread carriage of multiple strains of the same species. There is ample opportunity for between-species competition: the nasopharynx typically hosts tens or hundreds of species[16,17], while the gut typically hosts thousands[18]. The extent to which this extensive diversity may contribute to resistance evolution remains to be evaluated.

Alternative mechanisms that could explain coexistence between drug-sensitive and resistant pathogens have been proposed[4,6–8,38–40]. Some support only modest amounts of coexistence[4,6], while others may be less empirically generalisable, such as strongly age-assortative mixing[6,7], independent mappings of balancing selection[7], or specific immune responses to resistance-associated phenotypes[4,38–40]. We have focused on how within-host competition can promote substantial coexistence on its own. A more complex

model incorporating additional drivers of coexistence would support similar amounts of coexistence while diminishing the relative importance of within-host competition.

The models we have contrasted here make a number of simplifying assumptions. We have assumed that observed resistance patterns represent the equilibrium state, following from the lack of conclusive evidence for significant time lags in resistance prevalence[41] (Supplementary Note 6). We have assumed that antibiotics kill all sensitive cells instantaneously rather than adopting a more mechanistically-explicit model of treatment[42], and that host immunity completely prevents colonisation by previously-cleared serotypes rather than providing partial protection[27]. We have ignored effects of population structure, such as age-assortative mixing[6,7] and heterogeneity in antibiotic consumption[4,6], which may promote additional coexistence. We have assumed that co-colonisation occurs through sequential transmission, ignoring the alternative routes of *de novo* mutation (which may be especially important for long-lived chronic infections[43,44]), acquisition or loss of resistance through horizontal gene transfer, and simultaneous transmission of multiple strains from co-colonised carriers. Finally, we have focused on modelling resistance to a single drug at a time rather than exploring multi-drug resistance[45]. Elaborations of our simple mixed-carriage model incorporating these additional complexities may provide a means with which to explore the importance of these mechanisms.

Antibiotic resistance is one of the foremost threats to human health, and combating this threat will require the global deployment of coordinated interventions[9,10]. Mathematical models of disease transmission will play a crucial role in this endeavour, because they can explicitly integrate the mechanisms that drive resistance evolution in a population-level framework and allow us to quantify long-term trends as well as the likely impact and cost-effectiveness of any large-scale interventions for reducing resistance[46]. Providing a framework in which to answer public health questions demands a balance between mathematical tractability and necessary complexity; building on the simple model proposed here will help to establish that balance. If mathematical models incorporate a truly mechanistic understanding of resistance evolution, they will be better able to explain empirical patterns of resistance and accurately predict the impact of interventions at a national and global level[46].

## Methods

The problem of coexistence

*Data and sources* — We use data from the European Centre for Disease Prevention and Control (ECDC) on primary-care consumption of penicillins, fluoroquinolones, and macrolides[2] versus aminopenicillin resistance and fluoroquinolone resistance in *E. coli*, and macrolide non-susceptibility and penicillin non-susceptibility in *S. pneumoniae*[3], across up to 30 European countries. All data are from 2015, except for *S. pneumoniae* penicillin non-susceptibility versus penicillin consumption, which are from 2007 as breakpoints for *S. pneumoniae* penicillin non-susceptibility were changed in some countries after this year, yielding inconsistencies in resistance data between countries[4,47]. Antibiotic use is classified into primary-care and hospital consumption, with the majority of consumption in primary care[2]. We use primary-care data only, as we are focusing on community-acquired bacterial carriage. Resistance is measured from invasive isolates extracted from blood and cerebrospinal fluid[3]. We assume that each isolate is an unbiased sample of commensally-carried strains[7]. See Supplementary Note 7 for full details.

*Trends in resistance prevalence* — In Fig. 1a, linear regressions are least-squares fits to maximum-likelihood estimates of the resistance prevalence in each country. In Fig. 1d, the average resistance prevalence in Europe is calculated as the population-weighted mean of resistance prevalence across countries that reported data for all years in 2007–2015. See Supplementary Note 6 for more details.

Two models of within-host dynamics

In the Results section, we contrast two models of within-host dynamics: the existing knockout model[4,28], and the novel mixed-carriage model. Here, we describe how the knockout and mixed-carriage models can be implemented for two strains in a stochastic individual-based framework, then show how they can be approximated using systems of ODEs. The individual-based and ODE implementations are equivalent under certain limiting assumptions and produce similar results (Supplementary Note 1). We use the ODE implementations to illustrate coexistence between resistant and sensitive strains and for model fitting (Figs. 1, 3, & 4). The individual-based implementation of the mixed-carriage model can be extended to simulate an arbitrary number of strains (Supplementary Note 5) and is used to analyse serotype dynamics (Figs. 5 & 6).

*Knockout model*[4,28] — In a population of $N$ hosts indexed by $i \in [1..N]$, there are $N_{\mathrm{X}}$ non-carriers, $N_{\mathrm{S}}$ sensitive-strain carriers, $N_{\mathrm{R}}$ resistant-strain carriers, and $N_{\mathrm{SR}}$ dual carriers; we notate host $i$'s state as $h_i \in \{\mathrm{X}, \mathrm{S}, \mathrm{R}, \mathrm{SR}\}$. The following host-state transitions occur as inhomogeneous Poisson point processes at the specified per-host rates:

$$X \xrightarrow{\lambda_{\mathrm{s}}} S \quad (sensitive\ strain\ colonisation)$$
$$X \xrightarrow{\lambda_{\mathrm{r}}} R \quad (resistant\ strain\ colonisation)$$
$$S \xrightarrow{k\lambda_{\mathrm{r}}} SR \quad (sensitive\ strain\ co-colonisation)$$

$$R \xrightarrow{k\lambda_s} SR \quad (resistant\ strain\ co-colonisation)$$
$$SR \xrightarrow{\frac{1}{2}k\lambda_s} S \quad (knockout\ of\ resistant\ strain)$$
$$SR \xrightarrow{\frac{1}{2}k\lambda_r} R \quad (knockout\ of\ sensitive\ strain)$$
$$S \xrightarrow{u+\tau} X \quad (sensitive-strain\ carrier\ clearance\ or\ treatment)$$
$$R \xrightarrow{u} X \quad (resistant-strain\ carrier\ clearance)$$
$$SR \xrightarrow{u} X \quad (dual\ carrier\ clearance)$$
$$SR \xrightarrow{\tau} R \quad (dual\ carrier\ treatment)\ .$$

For example, non-carriers (X) become sensitive-strain carriers (S) at rate $\lambda_s$, and so on. Above, $\lambda_s = \beta \frac{N_S + \frac{1}{2}N_{SR}}{N}$ is the sensitive strain's force of infection, $\lambda_r = \beta(1-c)\frac{N_R + \frac{1}{2}N_{SR}}{N}$ is the resistant strain's force of infection, $\beta$ is the transmission rate, $c$ is the transmission cost of resistance, $k$ is the relative efficiency of co-colonisation, $u$ is the natural clearance rate, and $\tau$ is the treatment rate. In this model, the resistance prevalence is $\rho = (N_R + \frac{1}{2}N_{SR})/(N_S + N_R + N_{SR})$.

*Mixed-carriage model* — In a population of $N$ hosts indexed by $i \in [1..N]$ as above, host $i$'s state is $(s_i, r_i)$, where $s_i \geq 0$ is host $i$'s carriage of the sensitive strain and $r_i \geq 0$ is host $i$'s carriage of the resistant strain. In a non-carrier, $s_i = r_i = 0$, while in a carrier, $s_i + r_i = 1$. We model transmission, clearance, and treatment events as inhomogeneous Poisson point processes, while within-host strain growth is updated in each host at regular discrete time steps. The following host-state transitions occur at the specified per-host rates:

$$(s_i, r_i) \xrightarrow{\kappa_i\lambda_s} \left(\frac{s_i+\iota}{s_i+r_i+\iota}, \frac{r_i}{s_i+r_i+\iota}\right) \quad (sensitive\ strain\ transmission)$$
$$(s_i, r_i) \xrightarrow{\kappa_i\lambda_r} \left(\frac{s_i}{s_i+r_i+\iota}, \frac{r_i+\iota}{s_i+r_i+\iota}\right) \quad (resistant\ strain\ transmission)$$
$$(s_i, r_i) \xrightarrow{u} (0,0) \quad (clearance)$$
$$(s_i, r_i) \xrightarrow{\tau} \begin{cases} (0,0) & \text{if } r_i = 0 \\ (0,1) & \text{if } r_i > 0 \end{cases} \quad (treatment)\ .$$

For example, a host with state $(s_i, r_i) = (0,1)$ changes state to $(s_i, r_i) = (\frac{\iota}{1+\iota}, \frac{1}{1+\iota})$ at rate $\kappa_i\lambda_s$, and so on. Above, $\kappa_i = 1$ if $(s_i, r_i) = (0,0)$ and $\kappa_i = k$ otherwise; $\iota$ is the germ size; and force-of-infection terms are $\lambda_s = \beta \max(Y_{\min}, \sum_i s_i)/N$ and $\lambda_r = \beta(1-c)\max(Y_{\min}, \sum_i r_i)/N$, where we can set $Y_{\min} = 1$ to effectively assume there is always at least one carrier of each strain to avoid stochastic elimination of strains[27], or set $Y_{\min} = 0$ to not do this. The resistance prevalence is $\rho = \sum_i r_i / \sum_i (s_i + r_i)$.

Updates to within-host strain growth happen to all hosts simultaneously at intervals of $\Delta t$ (unless otherwise specified, $\Delta t$ = 0.001 mo$^{-1}$), as follows. For each host, any strains for which carriage is less than $f_{min}$ are set to zero (we primarily use $f_{min}$ = 3×10$^{-5}$ to keep strains from persisting when they reach low frequencies, but can set $f_{min}$ = 0 to allow them to remain at any frequency until treatment and/or natural clearance occurs). Then the sensitive strain in each carrier grows by a factor $\omega_s = w_s^{\Delta t}$, where $w_s$ is the sensitive strain's relative growth rate (such that $w_s$ = 1 translates to no differential within-host

growth). Finally, each colonised host's total carriage is normalised so that $s_i + r_i = 1$. That is, every $\Delta t$ units of time, each colonised host undergoes the transition

$$(s_i, r_i) \to \left( \frac{\omega_s q(s_i)}{\omega_s q(s_i) + q(r_i)}, \frac{q(r_i)}{\omega_s q(s_i) + q(r_i)} \right),$$

where

$$q(a) = \begin{cases} a & \text{if } a \geq f_{\min} \\ 0 & \text{if } a < f_{\min} \end{cases}.$$

In our implementation, we calculate the force-of-infection terms and the number of events of each type between time $t$ and $t + \Delta t$ during the "updating" step, then execute each event in a random order.

*Systems of ODEs* — The knockout and mixed-carriage models can be approximated using ODEs (Supplementary Note 1). Following previous work[4,28], the knockout model is implemented as

$$\begin{aligned} \frac{dS}{dt} &= \beta S_{\text{tot}} X - (u+\tau)S - k\beta(1-c)R_{\text{tot}}S + \frac{k\beta S_{\text{tot}} D}{2} \\ \frac{dR}{dt} &= \beta(1-c)R_{\text{tot}}X - uR - k\beta S_{\text{tot}}R + \frac{k\beta(1-c)R_{\text{tot}}D}{2} + \tau D \\ \frac{dD}{dt} &= k\beta(1-c)R_{\text{tot}}S + k\beta S_{\text{tot}}R - (u+\tau)D - \frac{k\beta S_{\text{tot}}D}{2} - \frac{k\beta(1-c)R_{\text{tot}}D}{2} \\ X &= 1 - S - R - D\,. \end{aligned} \tag{1}$$

Here, $S$ is the fraction of sensitive-strain carriers in the population; $R$ is the fraction of resistant-strain carriers; $D$ is the fraction of dual carriers (*i.e.*, SR hosts); and $X$ is the fraction of non-carriers. Here, $S_{\text{tot}} = S + D/2$ and $R_{\text{tot}} = R + D/2$ give the effective population burden of sensitive- and resistant-strain colonisation, respectively, and the resistance prevalence is $\rho = R_{\text{tot}}/(1–X)$. The parameters $\beta, c, u, \tau$, and $k$ correspond to those used in the individual-based implementation of the knockout model, described above.

Similarly, the mixed-carriage model (in the absence of differential within-host growth) can be approximated using the following system of ODEs:

$$\begin{aligned} \frac{dS}{dt} &= \beta S_{\text{tot}} X - (u+\tau)S - k\beta(1-c)R_{\text{tot}}S \\ \frac{dR}{dt} &= \beta(1-c)R_{\text{tot}}X - uR - k\beta S_{\text{tot}}R + \tau(S_R + R_S) \\ \frac{dS_R}{dt} &= k\beta(1-c)R_{\text{tot}}S - (u+\tau)S_R \\ \frac{dR_S}{dt} &= k\beta S_{\text{tot}}R - (u+\tau)R_S \\ X &= 1 - S - R - S_R - R_S\,. \end{aligned} \tag{2}$$

Here, the compartment $S_R$ captures the fraction of the population predominantly colonised with sensitive bacteria, but also carrying a small amount of resistant bacteria that are carried in insufficient quantity to transmit, and $S_{\text{tot}} = S + S_R$ gives the effective population burden of sensitive-strain colonisation. Similarly, the compartment $R_S$ captures the fraction of the population predominantly colonised with resistant bacteria, but also carrying a small amount of sensitive bacteria that are carried in insufficient

quantity to transmit, and $R_{\text{tot}} = R + R_S$ gives the effective population burden of resistant-strain colonisation. The overall resistance prevalence is $\rho = R_{\text{tot}}/(1–X)$. The parameters $\beta, c, u, \tau$, and $k$ correspond to those used in the individual-based implementation of the mixed-carriage model, described above.

Finally, the mixed-carriage model with differential within-host growth can be approximated with ODEs by adding "intermediate" compartments between $R_S$ and $S_R$:

$$\begin{aligned}
\frac{dS}{dt} &= \beta S_{\text{tot}} X - (u+\tau)S - k\beta(1-c)R_{\text{tot}}S + b_0 S_R \\
\frac{dR}{dt} &= \beta(1-c)R_{\text{tot}}X - uR + \tau\left(S_R + \sum_{v=1}^{Z} D_v + R_S\right) - k\beta S_{\text{tot}}R \\
\frac{dS_R}{dt} &= k\beta(1-c)R_{\text{tot}}S - (u+\tau)S_R - b_0 S_R + bD_1 \\
\frac{dD_v}{dt} &= -(u+\tau)D_v - bD_v + bD_{v+1} \qquad \text{for all } v \in [1..Z] \\
D_{Z+1} &\equiv R_S \\
\frac{dR_S}{dt} &= k\beta S_{\text{tot}}R - (u+\tau)R_S - bR_S \\
X &= 1 - S - R - S_R - \sum_{v=1}^{Z} D_v - R_S \,.
\end{aligned} \tag{3}$$

Here, there are $Z$ "intermediate" compartments between $R_S$ and $S_R$, labelled $D_1$ through $D_Z$ (we use $Z = 7$; see Supplementary Note 1 for a graphical illustration of the dynamics of these intermediate compartments). Here, $b$ determines the within-host growth rate of the sensitive strain relative to the resistant strain, setting the rate at which individuals move from the $R_S$ compartment through intermediate compartments and finally through to $S$ as the resistant strain is gradually outcompeted by the sensitive strain. A separate parameter $b_0$ sets the rate of the final transition from $S_R$ to $S$. In practice, we set $b_0 = \frac{1}{2}b$, which for $Z = 7$ and $\iota = 0.001$ corresponds to the resistant strain effectively becoming lost once its within-host frequency drops below $f_{\text{min}} = 3\times10^{-5}$ (Supplementary Note 1). The parameters $b$ and $b_0$ replace the parameters $w_s$ and $f_{\text{min}}$ from the individual-based implementation of the mixed-carriage model, above; all other parameters (*i.e.* $\beta, c, u, \tau$, and $k$) correspond to those used in the individual-based implementation.

Notating the fraction of a host's bacterial carriage that is resistant as $r_Y$ for a host with state $Y$, we assume that $r_{\text{R}_\text{S}} = \frac{1}{1+\iota}$, $r_{\text{S}_\text{R}} = \frac{\iota}{1+\iota}$, and that intermediate compartments are evenly spaced between these points on a logistic curve, *i.e.* that $r_{D_v} = \frac{1}{1+\exp(y(v))}$, where $y(v) = \log(\iota)\left(\frac{2v}{Z+1} - 1\right)$. We assume that individuals in compartment $D_v$ transmit the resistant strain a fraction $r_{D_v}$ of the time and transmit the sensitive strain a fraction $1 - r_{D_v}$ of the time, but that $R_S$ individuals only transmit the resistant strain and $S_R$ individuals only transmit the sensitive strain. Ignoring transmission of the minor strain for these two host types maintains consistency with equations (2) and maintains structural neutrality for equivalent strains in equations (3). Accordingly, in the model above, $S_{\text{tot}} = S + S_R + \sum_v D_v(1 - r_{D_v})$ and $R_{\text{tot}} = R + R_S + \sum_v D_v r_{D_v}$. Note that the

mixed-carriage model without differential within-host growth can be recovered from the above model by setting $b = b_0 = 0$; in model fitting, when we allow differential growth (*i.e.* $b > 0$) we assume that this accounts for the cost of resistance, and accordingly set $c = 0$. In this model, the overall resistance prevalence is $\rho = R_{\text{tot}}/(1-X)$.

*Initial conditions and solutions* — For all individual-based model simulations, we assume that 5% of hosts are colonised at the beginning of the simulation by a single randomly-selected strain, and run the simulation for 100–400 years, taking the average state over the last 50–100 years as the equilibrium state. Individual-based models are simulated in C++. All ODE models are solved by setting single-carriage compartments ($S$ and $R$) equal to 0.001 and all dual-carriage compartments to 0, then integrating the systems of ordinary differential equations numerically in C++ using the Runge–Kutta Dormand–Prince method until they reach equilibrium.

Model fitting to resistance prevalence in commensal bacteria

In the source data[2], antibiotic consumption rates are given in defined daily doses (DDD) per thousand people per day; we convert these to overall treatment rates by assuming that 10 DDD comprise one treatment course for penicillin[7] and fluoroquinolones, while 7 DDD comprise one treatment course for macrolides.

We use Bayesian inference to fit the model to empirical data, using differential evolution Markov chain Monte Carlo (DE-MCMC[48]) to estimate the posterior distribution of model parameters. We assume that the number of resistant isolates observed in a given country is binomially distributed; the probability of observing a resistant isolate is equal to the resistance prevalence $\rho$ predicted by the model, plus some additional dispersion modelled using a [0,1]–truncated normal distribution. Modelling the "true" resistance prevalence as a random variable allows us to account for between-country variation in resistance prevalence not captured by our dynamic model. As we assume that the only parameter that varies between European countries is the rate of antibiotic consumption, this additional variation is intended to account for other factors that may vary between countries, whether they are explicitly part of the model structure (*e.g.* transmission rates varying from country to country) or not (*e.g.* differences in laboratory procedures, population structure, or prescription patterns from country to country).

For a given model fit with parameters $\theta$, suppose that country $m$ (where countries are numbered 1 to $M$) has antibiotic treatment rate $\tau_m$ and reports that $r_m$ out of $n_m$ isolates are resistant. Over all $M$ countries, these data are denoted $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_M)$, $\boldsymbol{r} = (r_1, r_2, \dots, r_M)$, and $\boldsymbol{n} = (n_1, n_2, \dots, n_M)$, respectively. We also have $Y^{(0)}$ and $Y^{(1)}$, which are the lower and upper bounds for carriage prevalence in any country (see below). Together, $\boldsymbol{\tau}, \boldsymbol{r}, \boldsymbol{n}, Y^{(0)}$ and $Y^{(1)}$ are the data to which the model is being fit, and model parameters are $\theta = (\beta, c, b, u, k, \sigma)$. (Note that, for certain data sets, not all of the parameters in $\theta$ are permitted to vary; specifically, we assume $u = 1$ when fitting *S. pneumoniae* for consistency with previous studies, and we only allow one of $c$ and $b$ to vary at a time in order to contrast these two alternative costs of resistance.) Suppose that, for a given treatment rate $\tau_m$, the model predicts a resistance prevalence of $\rho(\tau_m|\theta)$ and a prevalence of carriage $Y(\tau_m|\theta)$. Then, the likelihood of the model fit is

$$\mathcal{L}\left(\boldsymbol{\tau}, \boldsymbol{r}, \boldsymbol{n}, Y^{(0)}, Y^{(1)}|\theta\right) = \prod_m \mathcal{C}\left(\tau_m, Y^{(0)}, Y^{(1)}|\theta\right) \mathcal{R}(\tau_m, r_m, n_m|\theta)\,,$$

which is constructed of two components that are evaluated for each country. The first component,

$$\mathcal{C}\left(\tau_m, Y^{(0)}, Y^{(1)}|\theta\right) = \begin{cases} 1 & \text{if } Y^{(0)} \leq Y(\tau_m|\theta) \leq Y^{(1)} \\ \exp\left(-1000\right) & \text{otherwise} \end{cases},$$

heavily penalises any model fit which predicts that any country has a prevalence of carriage not within the bounds $[Y^{(0)}, Y^{(1)}]$ and is used to prevent the model-fitting process from predicting an unrealistic carriage prevalence for any country. The second component,

$$\mathcal{R}(\tau_m, r_m, n_m|\theta) = \int_0^1 \mathcal{T}\left(x\middle|\mu = \rho(\tau_m|\theta), \sigma = \sigma(\theta)\right) \mathcal{B}(r_m|n = n_m, p = x)\,\mathrm{d}x\,,$$

assigns a likelihood to the model-predicted resistance prevalence $\rho(\tau_m|\theta)$ given that country $m$ has reported that $r_m$ of $n_m$ bacterial isolates are resistant. Above: $\mathcal{T}(x|\mu, \sigma) = \varphi(x|\mu, \sigma)/\left(\Phi(1|\mu, \sigma) - \Phi(0|\mu, \sigma)\right)$ is the probability density function (PDF) of a truncated normal distribution with bounds 0 and 1, where $\varphi(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ is the untruncated normal PDF and $\Phi(x|\mu, \sigma) = \frac{1}{2}(1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right))$ is the untruncated normal cumulative distribution function (CDF); and $\mathcal{B}(r|n, p) = \binom{n}{r} p^r (1-p)^{n-r}$ is the binomial distribution probability mass function (PMF), such that the integral calculates a weighted likelihood over all possible "true" resistance prevalences *x*. The parameter $\sigma(\theta)$ of the truncated normal distribution is fit as one of the parameters of the model so that between-country variation is estimated separately for each alternative model.

Priors used for model fitting, posterior distributions from model fitting, and further details of MCMC can be found in Supplementary Note 4. Note that since we are only fitting to the measured resistance prevalence in each European country and to a fixed range of carriage prevalence, the values of certain parameters are difficult to identify, particularly for the knockout model.

*Model comparison* — For each model fit, we calculate the Akaike Information Criterion $AIC = 2K - 2\log(\widehat{\mathcal{L}})$, where *K* is the number of free parameters and $\widehat{\mathcal{L}}$ is the maximum likelihood for a given model fit.

Patterns of resistance and coexistence among bacterial subtypes

For Figs. 5 & 6, we extend the individual-based mixed-carriage model to accommodate an arbitrary number of strains (Supplementary Note 5). For Fig. 5 only, we also introduce serotype-specific adaptive immunity. Hosts develop immunity to a serotype when they naturally clear that serotype, and immunity provides complete protection against future colonisations by that serotype. We assume that hosts are replaced by new, immunologically-naïve, uncolonised hosts at rate $\alpha$ = 1/60 mo$^{-1}$, reflecting the relative importance of hosts aged 5 years and under for pneumococcal transmission[4,49]. Other parameters for Fig. 5 are $\beta$ = 3.2 mo$^{-1}$ for sensitive strains and $\beta$ = 2.88 mo$^{-1}$ for

resistant strains (*i.e.* a 10% transmission cost of resistance), $w$ ranging from 1 to 30 for sensitive strains, where the serotype with the highest growth rate also has the longest duration of carriage, $w$ ranging from 0.8 to 24 for resistant strains (*i.e.* a 20% growth cost of resistance), $k = 1$, $\tau = 0.025$, and $N = 1\times10^6$. For Fig. 6, other parameters are $\beta = 2$ $mo^{-1}$, $u = 1$ $mo^{-1}$, $w = 1$, and $k = 1$ unless otherwise specified in the caption. The treatment rate is $\tau = 0$ for Fig. 6a and $\tau = 0.075$ for Fig. 6b for serotypes circulating both separately and in the same population. For Fig. 6c, the treatment rate is $\tau = 0.075$ when serotypes circulate together, but $\tau = 0.05$ when serotypes circulate individually. The reduced treatment rate when serotypes circulate individually is necessary to observe the trend in resistance prevalence among serotypes (with $\tau = 0.075$, all serotypes show 100% resistance prevalence, so trends are not apparent). We use a population size of $N = 1\times10^6$ for runs with serotypes circulating together, and $N = 2\times10^5$ for runs with serotypes circulating individually.

Code availability

C++ code for the individual-based model is available at https://github.com/nicholasdavies/tinyhost.

**Data availability**

All data used in this analysis are publicly available[2,3].

**Acknowledgements**

We thank M. Davies and A. Levy for assistance; S. Lehtinen, C. Colijn, and M. Lipsitch for discussion; and four anonymous reviewers for helpful comments. NGD, MJ and KEA were funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Immunisation at the London School of Hygiene and Tropical Medicine in partnership with Public Health England (PHE). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or PHE. For part of this work SF was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant number 208812/Z/17/Z).

**Author contributions**

NGD, SF, MJ and KEA conceived the study; NGD performed the analyses; NGD and KEA drafted the manuscript, which all authors revised.

**Competing interests**

The authors declare no competing financial interests.

**Supplementary Information**

Supplementary Notes 1–7 (with figures, tables, references, and appendices inline)

## References


1. Goossens, H., Ferech, M., Vander Stichele, R. & Elseviers, M. Outpatient antibiotic use in Europe and association with resistance: A cross-national database study. *Lancet* **365,** 579–587 (2005).
2. European Centre for Disease Prevention and Control. Antimicrobial consumption rates by country. (2018). Available at: http://ecdc.europa.eu/en/healthtopics/antimicrobial_resistance/esac-net-database/Pages/Antimicrobial-consumption-rates-by-country.aspx.
3. European Centre for Disease Prevention and Control. Data from the ECDC Surveillance Atlas - Antimicrobial resistance. (2016). Available at: https://ecdc.europa.eu/en/antimicrobial-resistance/surveillance-and-disease-data/data-ecdc. (Accessed: 24th February 2018)
4. Colijn, C. *et al.* What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae*? *J. R. Soc. Interface* **7,** 905–919 (2010).
5. Hardin, G. The competitive exclusion principle. *Science (80-. ).* **131,** 1292–1297 (1960).
6. Cobey, S. *et al.* Host population structure and treatment frequency maintain balancing selection on drug resistance. *J. R. Soc. Interface* **14,** 20170295 (2017).
7. 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).
8. Austin, D. J., Kristinsson, K. G. & Anderson, R. M. The relationship between the volume of antimicrobial consumption in human communities and the frequency of resistance. *Proc. Natl. Acad. Sci.* **96,** 1152–1156 (1999).
9. World Health Organization. United Nations high-level meeting on antimicrobial resistance. (2016). Available at: http://www.who.int/mediacentre/events/2016/antimicrobial-resistance/en/%5Cnhttp://www.who.int/antimicrobial-resistance/events/UNGA-meeting-amr-sept2016/en/.
10. O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. (2016). Available at: https://amr-review.org/.
11. Kamng'ona, A. W. *et al.* High multiple carriage and emergence of *Streptococcus pneumoniae* vaccine serotype variants in Malawian children. *BMC Infect. Dis.* **15,** 234 (2015).
12. Turner, P. *et al.* Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *J. Clin. Microbiol.* **49,** 1784–1789 (2011).
13. Martinez-Medina, M. *et al.* Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* **15,** 872–882 (2009).
14. Mongkolrattanothai, K. *et al.* Simultaneous carriage of multiple genotypes of *Staphylococcus aureus* in children. *J. Med. Microbiol.* **60,** 317–322 (2011).
15. Gordon, D. M., O'Brien, C. L. & Pavli, P. Escherichia coli diversity in the lower intestinal tract of humans. *Environ. Microbiol. Rep.* **7,** 642–648 (2015).
16. Chaban, B. *et al.* Characterization of the Upper Respiratory Tract Microbiomes of

Patients with Pandemic H1N1 Influenza. *PLoS One* **8,** 1–11 (2013).
17. Ederveen, T. H. A. *et al.* Haemophilus is overrepresented in the nasopharynx of infants hospitalized with RSV infection and associated with increased viral load and enhanced mucosal CXCL8 responses. *Microbiome* **6,** 1–13 (2018).
18. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489,** 220–230 (2012).
19. Negri, M. C., Lipsitch, M., Blázquez, J., Levin, B. R. & Baquero, F. Concentration-dependent selection of small phenotypic differences in TEM beta-lactamase-mediated antibiotic resistance. *Antimicrob. Agents Chemother.* **44,** 2485–91 (2000).
20. Wargo, A. R., Huijben, S., de Roode, J. C., Shepherd, J. & Read, A. F. Competitive release and facilitation of drug-resistant parasites after therapeutic chemotherapy in a rodent malaria model. *Proc. Natl. Acad. Sci.* **104,** 19914–19919 (2007).
21. Melnyk, A. H., Wong, A. & Kassen, R. The fitness costs of antibiotic resistance mutations. *Evol. Appl.* **8,** 273–283 (2015).
22. Smani, Y. *et al.* In vitro and in vivo reduced fitness and virulence in ciprofloxacin-resistant Acinetobacter baumannii. *Clin. Microbiol. Infect.* **18,** 1–4 (2012).
23. Birch, L. C. The meanings of competition. *Am. Nat.* **91,** 5–18 (1957).
24. Hastings, I. M. Complex dynamics and stability of resistance to antimalarial drugs. *Parasitology* **132,** 615–624 (2006).
25. Ayala, F. J. Competition between species: frequency dependence. *Science (80-. ).* **171,** 820–824 (1971).
26. Ayala, F. J. & Campbell, C. A. Frequency-dependent selection. *Annu. Rev. Ecol. Syst.* **5,** 115–138 (1974).
27. Cobey, S. & Lipsitch, M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science (80-. ).* **335,** 1376–1380 (2012).
28. 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).
29. Sinervo, B. & Lively, C. M. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380,** 240–243 (1996).
30. Gigord, L. D. B., Macnair, M. R. & Smithson, A. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid *Dactylorhiza sambucina* (L.) Soò. *Proc. Natl. Acad. Sci.* **98,** 6253–6255 (2001).
31. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394,** 69–72 (1998).
32. Wale, N. *et al.* Resource limitation prevents the emergence of drug resistance by intensifying within-host competition. *Proc. Natl. Acad. Sci.* 201715874 (2017). doi:10.1073/pnas.1715874115
33. Lewnard, J. A. *et al.* Impact of antimicrobial treatment for acute otitis media on carriage dynamics of penicillin-susceptible and penicillin–non-susceptible Streptococcus pneumoniae. *J. Infect. Dis.* jiy343 (2018). doi:10.1093/infdis/jiy343/5033369
34. Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? *Current Opinion in Microbiology* **9,** 461–465 (2006).

35. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: Is it possible to reverse resistance? *Nat. Rev. Microbiol.* **8,** 260–271 (2010).
36. Flasche, S. *et al.* The impact of specific and non-specific immunity on the ecology of *Streptococcus pneumoniae* and the implications for vaccination. *Proc. R. Soc. B Biol. Sci.* **280,** 20131939 (2013).
37. MacFadden, D. R., McGough, S. F., Fisman, D., Santillana, M. & Brownstein, J. S. Antibiotic resistance increases with local temperature. *Nat. Clim. Chang.* **8,** 510–514 (2018).
38. Dietz, K. Epidemiologic interference of virus populations. *J. Math. Biol.* **8,** 291–300 (1979).
39. Gupta, S., Swinton, J. & Anderson, R. M. Theoretical studies of the effects of heterogeneity in the parasite population on the transmission dynamics of malaria. *Proc. R. Soc. B Biol. Sci.* **256,** 231–238 (1994).
40. Lipsitch, M. Vaccination against colonizing bacteria with multiple serotypes. *Proc. Natl. Acad. Sci.* **94,** 6571–6576 (1997).
41. Blanquart, F., Lehtinen, S. & Fraser, C. An evolutionary model to predict the frequency of antibiotic resistance under seasonal antibiotic use, and an application to *Streptococcus pneumoniae*. *Proc. R. Soc. B Biol. Sci.* **284,** 20170679 (2017).
42. Colijn, C. & Cohen, T. How competition governs whether moderate or aggressive treatment minimizes antibiotic resistance. *Elife* **4,** 1–29 (2015).
43. Smith, E. E. *et al.* Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl. Acad. Sci.* **103,** 8487–8492 (2006).
44. Yang, L. *et al.* Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl. Acad. Sci.* **108,** 7481–7486 (2011).
45. Lehtinen, S. *et al.* Mechanisms that maintain coexistence of antibiotic sensitivity and resistance also promote high frequencies of multidrug resistance. *bioRxiv* **14/12/17,** 1–17 (2017).
46. Atkins, K. E. *et al.* Use of mathematical modelling to assess the impact of vaccines on antibiotic resistance. *Lancet Infect. Dis.* (2017). doi:10.1016/S1473-3099(17)30478-4
47. Goossens, M. C., Catry, B. & Verhaegen, J. Antimicrobial resistance to benzylpenicillin in invasive pneumococcal disease in Belgium, 2003-2010: The effect of altering clinical breakpoints. *Epidemiol. Infect.* **141,** 490–495 (2013).
48. Ter Braak, C. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces. *Stat. Comput.* **16,** 239–249 (2006).
49. Bogaert, D. *et al.* Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *Lancet* **363,** 1871–1872 (2004).
50. Chewapreecha, C. *et al.* Dense genomic sampling identifies highways of pneumococcal recombination. *Nat. Genet.* **46,** 305–309 (2014).

*Page and reference numbers in the Supplementary Notes are separate from those of the main text.*

Supplementary information for

# Within-host dynamics shape antibiotic resistance in commensal bacteria

Nicholas G. Davies[1,2*], Stefan Flasche[1,2], Mark Jit[1,2,3], Katherine E. Atkins[1,2,4]

[1] Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.
[2] Department for Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.
[3] Modelling and Economics Unit, Public Health England, London SE1 8UG, UK.
[4] Centre for Global Health, Usher Institute of Population Health Sciences and Informatics, Edinburgh Medical School, The University of Edinburgh, Edinburgh EH8 9AG, UK.
* To whom correspondence should be addressed. E-mail: Nicholas.Davies@lshtm.ac.uk

## Contents

## Supplementary Note 1. The stochastic individual-based and ODE implementations of the mixed-carriage model are equivalent

*Overview — In the main text, we introduce the "mixed-carriage" model for analysing the evolution of antibiotic resistance in commensal bacteria. We provide details of two alternative implementations of this model: one which uses stochastic individual-based methods and one which uses deterministic systems of ordinary differential equations (ODEs). We have asserted that the two implementations are equivalent under certain limiting assumptions, and in this supplementary note we provide some evidence for that assertion.*

*Section 1.1 formally shows that, under certain limiting assumptions, the individual-based and ODE implementations of the mixed-carriage model are equivalent in the absence of within-host strain growth.*

*Section 1.2 gives details on how within-host strain growth is introduced to the ODE implementation by approximating growth using a series of discrete steps.*

*Section 1.3 discusses how within-host competitive exclusion—that is, cells of one strain being eliminated completely from a host on account of being "crowded out" by other strains—is implemented in both the individual-based and ODE implementations.*

*Section 1.4 shows graphically that the individual-based and ODE implementations produce similar results.*

### 1.1 Equivalence of the individual-based and ODE implementations in the absence of within-host strain growth

In this section, we formally show that when $f_{\min} = Y_{\min} = 0$, $w_s = 1$, $\iota$ and $\Delta t$ are arbitrarily close to zero, and the population size $N$ is infinite, the individual-based and ODE implementations of the mixed-carriage model are equivalent. Differential within-host strain growth is discussed in section 1.2.

Briefly, the equivalence of the individual-based and ODE implementations can be seen by interpreting the rates of change in the system of ODEs described by equation (2) in the main text as rates of transitions between host states, verifying that these transition rates are equivalent to the event rates used in the individual-based implementation, and noting that events have an equivalent impact upon hosts in the individual-based implementation as the transitions in the ODE implementation do.

Suppose that, in the individual-based mixed-carriage model implementation, we have: two strains, no minimum host carriage frequency ($f_{\min} = 0$), no minimum number of carriers of each strain ($Y_{\min} = 0$), the germ size $\iota$ infinitesimally small, and equal within-host fitness for both strains ($w_s = 1$). Recall that we denote host $i$ as $h_i = (s_i, r_i)$, where $s_i$ is the host's sensitive-strain carriage and $r_i$ is the host's resistant-strain carriage. Suppose further that at some time $t$ there are $N$ hosts in total, and that of these $N$ hosts, $N_X$ hosts are non-carriers (*i.e.* $N_X$ hosts have host state $h_i = (0,0)$), $N_S$ hosts carry only

the sensitive strain ($h_i = (1,0)$), $N_{\mathrm{R}}$ hosts carry only the resistant strain ($h_i = (0,1)$), $N_{\mathrm{S_R}}$ hosts carry the sensitive strain plus a very small amount of the resistant strain ($h_i = (1-\delta_i, \delta_i)$,where all $\delta_i$ are infinitesimally close to zero), and $N_{\mathrm{R_S}}$ hosts carry the resistant strain plus a very small amount of the sensitive strain ($h_i = (\delta_i, 1-\delta_i)$). We are assuming that, at time $t$, all hosts can be classified as one of these five host types, so $N = N_{\mathrm{X}} + N_{\mathrm{S}} + N_{\mathrm{R}} + N_{\mathrm{S_R}} + N_{\mathrm{R_S}}$. Since all $\delta_i$ terms are infinitesimally small and $Y_{\min} = 0$, the force of infection terms $\lambda_{\mathrm{s}} = \beta \max(Y_{\min}, \sum_i s_i)/N$ and $\lambda_{\mathrm{r}} = \beta(1-c)\max(Y_{\min}, \sum_i r_i)/N$ can be simply written $\lambda_{\mathrm{s}} = \frac{\beta(N_{\mathrm{S}}+N_{\mathrm{S_R}})}{N}$ and $\lambda_{\mathrm{r}} = \frac{\beta(1-c)(N_{\mathrm{R}}+N_{\mathrm{R_S}})}{N}$.

The individual-based model implementation proceeds via (i) events of transmission, clearance, and treatment modelled as inhomogeneous Poisson point processes, and (ii) updates to within-host strain growth, which occur regularly at time intervals of $\Delta t$. Recall that the updating step applies the transition

$$(s_i, r_i) \to \left(\frac{\omega_{\mathrm{s}} q(s_i)}{\omega_{\mathrm{s}} q(s_i) + q(r_i)}, \frac{q(r_i)}{\omega_{\mathrm{s}} q(s_i) + q(r_i)}\right),$$

to all hosts with non-zero carriage, where

$$q(a) = \begin{cases} a & \text{if } a \geq f_{\min} \\ 0 & \text{if } a < f_{\min} \end{cases}$$

and $\omega_{\mathrm{s}} = w_{\mathrm{s}}^{\Delta t}$. Note that when $w_{\mathrm{s}} = 1$ and $f_{\min} = 0$, this updating step has no effect, so it can be ignored for our purposes.

Recall that the "events" in the mixed-carriage model are

$$\begin{aligned}
&(s_i, r_i) \xrightarrow{\kappa_i \lambda_{\mathrm{s}}} \left(\tfrac{s_i+\iota}{s_i+r_i+\iota}, \tfrac{r_i}{s_i+r_i+\iota}\right) \quad (sensitive\ strain\ transmission)\\
&(s_i, r_i) \xrightarrow{\kappa_i \lambda_{\mathrm{r}}} \left(\tfrac{s_i}{s_i+r_i+\iota}, \tfrac{r_i+\iota}{s_i+r_i+\iota}\right) \quad (resistant\ strain\ transmission)\\
&(s_i, r_i) \xrightarrow{u} (0,0) \quad (clearance)\\
&(s_i, r_i) \xrightarrow{\tau} \begin{cases} (0,0) & \text{if } r_i = 0 \\ (0,1) & \text{if } r_i > 0 \end{cases} \quad (treatment)\,.
\end{aligned}$$

Recalling that $\kappa_i = 1$ if $(s_i, r_i) = (0,0)$ and $\kappa_i = k$ otherwise, we can write out these transitions for each of the five host types, yielding:

*sensitive strain transmission*

"X" $(0,0) \xrightarrow{\lambda_{\mathrm{s}}} (1,0)$ "S"

"S" $(1,0) \xrightarrow{k\lambda_{\mathrm{s}}} (1,0)$ "S"

"R" $(0,1) \xrightarrow{k\lambda_{\mathrm{s}}} \left(\frac{\iota}{1+\iota}, \frac{1}{1+\iota}\right) \Leftrightarrow (\delta_i', 1-\delta_i')$ "$\mathrm{R_S}$"

"$\mathrm{S_R}$" $(1-\delta_i, \delta_i) \xrightarrow{k\lambda_{\mathrm{s}}} \left(\frac{1-\delta_i+\iota}{1+\iota}, \frac{\delta_i}{1+\iota}\right) \Leftrightarrow (1-\delta_i', \delta_i')$ "$\mathrm{S_R}$"

"$\mathrm{R_S}$" $(\delta_i, 1-\delta_i) \xrightarrow{k\lambda_{\mathrm{s}}} \left(\frac{\delta_i+\iota}{1+\iota}, \frac{1-\delta_i}{1+\iota}\right) \Leftrightarrow (\delta_i', 1-\delta_i')$ "$\mathrm{R_S}$"

*resistant strain transmission*

$$\text{“X” } (0,0) \xrightarrow{\lambda_\text{r}} (0,1) \text{ “R”}$$
$$\text{“S” } (1,0) \xrightarrow{k\lambda_\text{r}} \left(\tfrac{1}{1+\iota}, \tfrac{\iota}{1+\iota}\right) \Leftrightarrow (1-\delta_i', \delta_i') \text{ “S}_\text{R}\text{”}$$
$$\text{“R” } (0,1) \xrightarrow{k\lambda_\text{r}} (0,1) \text{ “R”}$$
$$\text{“S}_\text{R}\text{” } (1-\delta_i, \delta_i) \xrightarrow{k\lambda_\text{r}} \left(\tfrac{1-\delta_i}{1+\iota}, \tfrac{\delta_i+\iota}{1+\iota}\right) \Leftrightarrow (1-\delta_i', \delta_i') \text{ “S}_\text{R}\text{”}$$
$$\text{“R}_\text{S}\text{” } (\delta_i, 1-\delta_i) \xrightarrow{k\lambda_\text{r}} \left(\tfrac{\delta_i}{1+\iota}, \tfrac{1-\delta_i+\iota}{1+\iota}\right) \Leftrightarrow (\delta_i', 1-\delta_i') \text{ “R}_\text{S}\text{”}$$

*clearance*

$$\text{“X” } (0,0) \xrightarrow{u} (0,0) \text{ “X”}$$
$$\text{“S” } (1,0) \xrightarrow{u} (0,0) \text{ “X”}$$
$$\text{“R” } (0,1) \xrightarrow{u} (0,0) \text{ “X”}$$
$$\text{“S}_\text{R}\text{” } (1-\delta_i, \delta_i) \xrightarrow{u} (0,0) \text{ “X”}$$
$$\text{“R}_\text{S}\text{” } (\delta_i, 1-\delta_i) \xrightarrow{u} (0,0) \text{ “X”}$$

*treatment*

$$\text{“X” } (0,0) \xrightarrow{\tau} (0,0) \text{ “X”}$$
$$\text{“S” } (1,0) \xrightarrow{\tau} (0,0) \text{ “X”}$$
$$\text{“R” } (0,1) \xrightarrow{\tau} (0,1) \text{ “R”}$$
$$\text{“S}_\text{R}\text{” } (1-\delta_i, \delta_i) \xrightarrow{\tau} (0,1) \text{ “R”}$$
$$\text{“R}_\text{S}\text{” } (\delta_i, 1-\delta_i) \xrightarrow{\tau} (0,1) \text{ “R”} \qquad \text{(S1)}$$

where some of the states on the right-hand side of each transition have been rewritten using $\delta_i'$, which are arbitrary values infinitesimally close to zero but which may differ from $\delta_i$. Since all $\delta_i$ and $\delta_i'$ are infinitesimally small, their precise values have no impact upon the overall model dynamics. In the transitions above, we have been able to classify all potential host states *after* events occur as one of the original five host states, so these five host states are sufficient to capture the full dynamics of the individual-based model.

Finally, it is a property of the Poisson distribution that when $X_i \sim \text{Poisson}(x_i)$ for all $i$, $\sum_i X_i \sim \text{Poisson}(\sum_i x_i)$. In other words, an event which happens at rate $x$ to individual hosts of type A will happen at rate $N_A x$ to all hosts of type A collectively.

Taking this all together, over a sufficiently small period of time $\Delta t$, such that only one event occurs within the period, transitions (S1) will have the following impact upon the number of hosts of each type:

$$\begin{aligned}
\Delta N_\text{S} &= T_{\text{X}\to\text{S}} - T_{\text{S}\to\text{X}} - T_{\text{S}\to\text{S}_\text{R}} \\
\Delta N_\text{R} &= T_{\text{X}\to\text{R}} - T_{\text{R}\to\text{X}} - T_{\text{R}\to\text{R}_\text{S}} + T_{\text{S}_\text{R}\to\text{R}} + T_{\text{R}_\text{S}\to\text{R}} \\
\Delta N_{\text{S}_\text{R}} &= T_{\text{S}\to\text{S}_\text{R}} - T_{\text{S}_\text{R}\to\text{X}} - T_{\text{S}_\text{R}\to\text{R}} \\
\Delta N_{\text{R}_\text{S}} &= T_{\text{R}\to\text{R}_\text{S}} - T_{\text{R}_\text{S}\to\text{X}} - T_{\text{R}_\text{S}\to\text{R}} \\
\Delta N_\text{X} &= -\Delta N_\text{S} - \Delta N_\text{R} - \Delta N_{\text{S}_\text{R}} - \Delta N_{\text{R}_\text{S}} \,, \qquad \text{(S2)}
\end{aligned}$$

where

$$
\begin{aligned}
T_{\mathrm{X}\to\mathrm{S}} &\sim \mathrm{Poisson}\left(\frac{\beta\left(N_{\mathrm{S}}+N_{\mathrm{S_R}}\right)}{N}N_{\mathrm{X}}\Delta t\right)\\
T_{\mathrm{S}\to\mathrm{X}} &\sim \mathrm{Poisson}\left((u+\tau)N_{\mathrm{S}}\Delta t\right)\\
T_{\mathrm{S}\to\mathrm{S_R}} &\sim \mathrm{Poisson}\left(k\frac{\beta(1-c)\left(N_{\mathrm{R}}+N_{\mathrm{R_S}}\right)}{N}N_{\mathrm{S}}\Delta t\right)\\
T_{\mathrm{X}\to\mathrm{R}} &\sim \mathrm{Poisson}\left(\frac{\beta(1-c)\left(N_{\mathrm{R}}+N_{\mathrm{R_S}}\right)}{N}N_{\mathrm{X}}\Delta t\right)\\
T_{\mathrm{R}\to\mathrm{X}} &\sim \mathrm{Poisson}(uN_R\Delta t)\\
T_{\mathrm{R}\to\mathrm{R_S}} &\sim \mathrm{Poisson}\left(k\frac{\beta\left(N_{\mathrm{S}}+N_{\mathrm{S_R}}\right)}{N}N_{\mathrm{R}}\Delta t\right)\\
T_{\mathrm{S_R}\to\mathrm{R}} &\sim \mathrm{Poisson}\left(\tau N_{\mathrm{S_R}}\Delta t\right)\\
T_{\mathrm{R_S}\to\mathrm{R}} &\sim \mathrm{Poisson}\left(\tau N_{\mathrm{R_S}}\Delta t\right)\\
T_{\mathrm{S_R}\to\mathrm{X}} &\sim \mathrm{Poisson}\left(uN_{\mathrm{S_R}}\Delta t\right)\\
T_{\mathrm{R_S}\to\mathrm{X}} &\sim \mathrm{Poisson}\left(uN_{\mathrm{R_S}}\Delta t\right).
\end{aligned}
\tag{S3}
$$

This is the stochastic, finite-population, individual-based analogue of the deterministic, infinite-population, ODE-based mixed-carriage model. To see this, substitute variates (S3) into equations (S2), divide both sides by $N\Delta t$, and make a change of variables such that $S=\frac{N_{\mathrm{S}}}{N}, R=\frac{N_{\mathrm{R}}}{N}, S_R=\frac{N_{\mathrm{S_R}}}{N}, R_S=\frac{N_{\mathrm{R_S}}}{N}$, and $X=\frac{N_{\mathrm{X}}}{N}$. Also, allow the population size $N$ to go to infinity, which permits replacing all variates of the form $T_{A\to B}\sim\mathrm{Poisson}(\alpha_{A\to B}\Delta t)$, where A and B are any two host types, with their expected values, $\mathrm{E}(T_{A\to B})=\alpha_{A\to B}\Delta t$. (For example, replace $T_{\mathrm{R_S}\to\mathrm{X}}\sim\mathrm{Poisson}\left(uN_{\mathrm{R_S}}\Delta t\right)$ with $\mathrm{E}\left(T_{\mathrm{R_S}\to\mathrm{X}}\right)=uN_{\mathrm{R_S}}\Delta t$.) This yields

$$
\begin{aligned}
\frac{\Delta S}{\Delta t} &= \beta(S+S_R)X-(u+\tau)S-k\beta(1-c)(R+R_S)S\\
\frac{\Delta R}{\Delta t} &= \beta(1-c)(R+R_S)X-uR-k\beta(S+S_R)R+\tau(S_R+R_S)\\
\frac{\Delta S_R}{\Delta t} &= k\beta(1-c)(R+R_S)S-(u+\tau)S_R\\
\frac{\Delta R_S}{\Delta t} &= k\beta(S+S_R)R-(u+\tau)R_S\\
X &= 1-S-R-S_R-R_S\,.
\end{aligned}
$$

By taking the limit as $\Delta t\to 0$, this gives the mixed-carriage ODE model implementation,

$$
\begin{aligned}
\frac{dS}{dt} &= \beta S_{\mathrm{tot}}X-(u+\tau)S-k\beta(1-c)R_{\mathrm{tot}}S\\
\frac{dR}{dt} &= \beta(1-c)R_{\mathrm{tot}}X-uR-k\beta S_{\mathrm{tot}}R+\tau(S_R+R_S)\\
\frac{dS_R}{dt} &= k\beta(1-c)R_{\mathrm{tot}}S-(u+\tau)S_R\\
\frac{dR_S}{dt} &= k\beta S_{\mathrm{tot}}R-(u+\tau)R_S\\
X &= 1-S-R-S_R-R_S\,.
\end{aligned}
$$

Therefore, the individual-based and ODE implementations are equivalent under the stipulated limiting assumptions. The equivalence of the two implementations of the

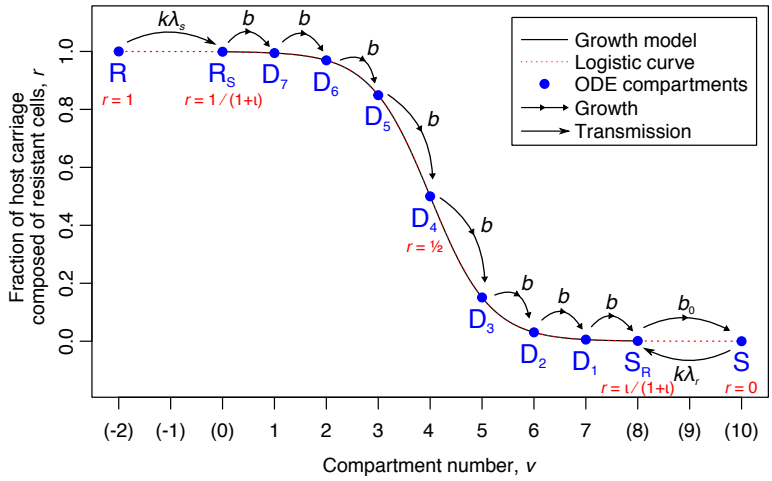



**Fig. S1 | Schematic of within-host strain growth.** Carriage compartments (filled blue circles) in the ODE approximation of the mixed-carriage model with differential within-host strain growth. Compartments between $R_S$ and $S_R$ are equally spaced along a logistic curve (dashed red line). The "growth model" curve (solid black line) was generated by the individual-based model and also defines a logistic curve. The transition from R to $R_S$ occurs via transmission of the sensitive strain at rate $k\lambda_s$, transitions between compartments $R_S$, $D_Z$, $D_{Z-1}$, ..., $D_1$, and $S_R$ occur via within-host strain growth at rate $b$, and the final transition from $S_R$ to S occurs at rate $b_0$. Additionally, a transition from S "back" to $S_R$ may occur via transmission of the resistant strain at rate $k\lambda_r$.

---

knockout model can be seen in a similar way, noting that each transition specified by the individual-based implementation corresponds to a term in the ODE implementation, which makes the implementations equivalent when the population size $N$ is very large.

### 1.2 Within-host strain growth in the individual-based and ODE implementations

The individual-based implementation captures differential within-host strain growth by eliminating any strains with a host carriage frequency of less than $f_{\min}$, multiplying the size of the sensitive strain by a factor $\omega_s = w_s^{\Delta t}$ —leaving the size of the resistant strain unchanged—then normalising each carrier's overall strain carriage so that the size of the sensitive strain and the size of the resistant strain sum to 1; this procedure is carried out for each host at time steps separated by $\Delta t$. Because of the normalisation step, there is only one degree of freedom in the system, and the size of the sensitive strain relative to the resistant strain follows a predictable curve. Technically, additional co-colonisation events will slightly speed up or slow down the movement along this curve, but as an approximation, we can ignore this effect on the grounds that it will not change dynamics of strain growth very much. Accordingly, to capture strain growth in our ODE-based implementation of the mixed-carriage model, we simulate within-host growth by moving dual carriers along discrete points on this curve, as illustrated in **Fig. S1**.

Note that the size of an exponentially-growing strain, relative to the combined size of itself and another exponentially-growing or non-growing strain, can be written $\frac{\exp(\alpha t)}{1+\exp(\alpha t)}$, which defines a logistic curve. Accordingly, in the ODE implementation, we assume that strain frequencies follow a logistic curve over time, with dual carriers moving between discrete points along this curve at rate $b$. **Fig. S1** illustrates this movement, and also

shows that the logistic curve used by the ODE implementation is indistinguishable from the explicit growth model used by the individual-based implementation. Note that as the number of intermediate compartments $Z$ approaches infinity, the relationship between $b$ and $w_s$ is approximately $b = -(Z+1)\frac{\log w_s}{\log \iota}$.

### 1.3 Within-host competitive exclusion

Because we assume within-host strain growth is exponential, it would be technically possible for the resistant strain to be driven to lower and lower within-host frequencies by the growth of the sensitive strain, and yet never reach zero frequency. This might be undesirable, as it could result in a situation where antibiotic treatment eliminates the sensitive strain from carriage and allows the resistant strain to completely take over the host in spite of the within-host frequency of the resistant strain being extremely low—possibly so low that it would correspond to less than a single cell. This could unfairly promote coexistence, because it would effectively allow the frequency-dependent advantage of resistant strains to remain the same regardless of the relative growth rate of the sensitive strain. To avoid this unrealistic scenario, we stipulate that strains below a certain within-host frequency are eliminated completely. In the individual-based model implementation, this is done using the parameter $f_{\min}$—any strain whose within-host frequency falls below this value is eliminated during the host "updating" step (see Methods). To control this behaviour in the ODE model implementation, we use the parameter $b_0$, which determines how quickly resistant cells are eliminated from $S_R$ carriers (because it is the rate of the transition from host state $S_R$ to host state S). In the individual-based mixed-carriage model, we assume that $f_{\min} = 3\times10^{-5}$, which means that strains are eliminated once they reach 3% of the germ size, $\iota = 0.001$.

In order to match this behaviour in the ODE implementation, we set $b_0 = \frac{1}{2}b$. This corresponds approximately to $f_{\min} = 3\times10^{-5}$ for the following reasons. First, recall (Methods) that the proportion of a host's carried cells that are resistant

$$r_{D_v} = \frac{1}{1+\exp\big(y(v)\big)},$$

where $y(v) = \log(\iota)\left(\frac{2v}{Z+1} - 1\right)$, $v$ is the ODE compartment number (**Fig. S1**), and $Z$ is the number of intermediate compartments. Note that we have $r_{D_v} = 3.16\times10^{-5}$ when $v = -2$, $\iota = 0.001$, and $Z = 7$ (we assume $\iota = 0.001$ and $Z = 7$ throughout the paper). That is, starting from the $S_R$ compartment (equivalent to compartment $v = 0$) on **Fig. S1**, if we assume it takes two additional "growth steps" to the right in order to completely eliminate the resistant strain and reach the S compartment, this takes us to a resistant-strain frequency of about $r_{D_v} = 3\times10^{-5}$. It takes twice as long to take two steps as it does to take one step, which is why we assume the rate of this $S_R$ to S transition is equal to half the normal rate, *i.e.* $b_0 = \frac{1}{2}b$.

This still leaves the question open as to whether our chosen value of $f_{\min} = 3\times10^{-5}$ is realistic. However, the minimum infective dose for *S. pneumoniae* has been estimated to

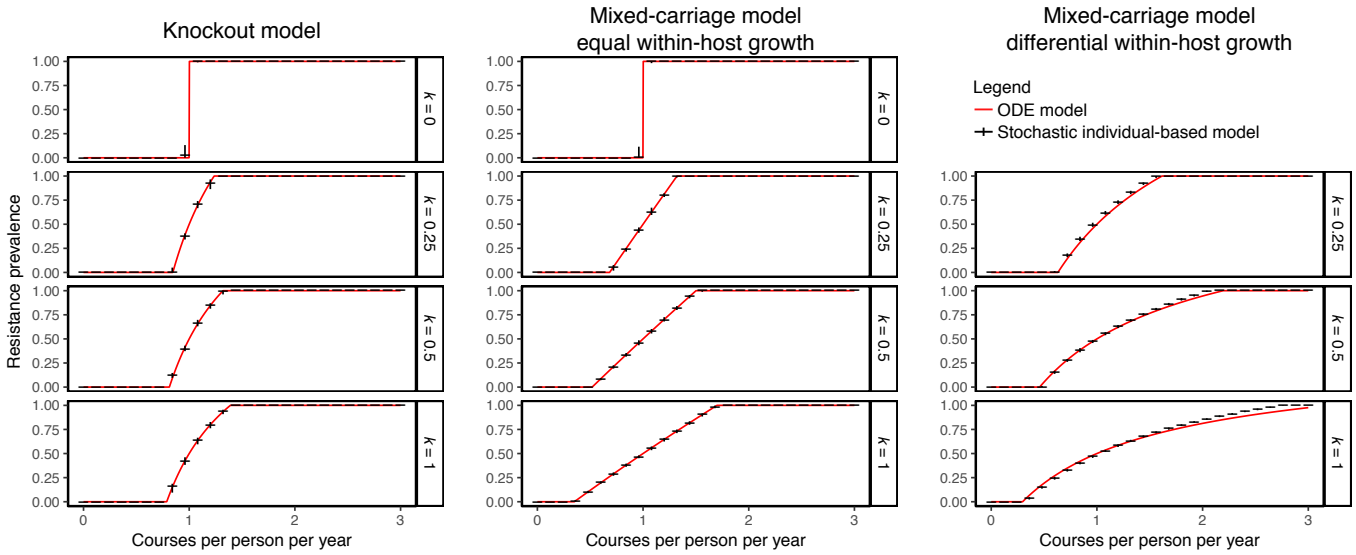



**Fig. S2 | Correspondence between ODE and individual-based model implementations.** We recapitulate the right-hand column of Fig. 3 of the main text to show that the individual-based implementations (black crosses) produce similar results to the ODE implementations (red lines) of each model. We assume $\iota = 0.001$, $f_{\min} = 3{\times}10^{-5}$, and $N$ = 100,000. Across all models, we have $\beta = 5$ mo$^{-1}$, $u$ = 1 mo$^{-1}$, and $k$ and $\tau$ as shown. Values of $c$, $b$, and $w_s$ are chosen so that resistance prevalence passes through 0.5 at $\tau = 1$ y$^{-1}$. Specifically, for the knockout model, from top to bottom, we have $c$ = 0.0769, 0.0660, 0.0538, and 0.0383; for the mixed-carriage model with equal growth, we have $c$ = 0.0769, 0.0992, 0.1112, and 0.1241; and for the mixed-carriage model with differential growth, we have $b$ = 3.8225, 3.1698, and 2.8644 for the ODE implementation with $Z$ = 7 intermediate compartments, and $w_s$ = 34, 20, and 14 for the individual-based implementation ($w_s$ values were estimated rather than chosen with an automated model calibration procedure). In all cases, the individual-based models were run for 200 years, with the vertical bars of each black cross showing the 95% interquantile range over the last 150 years of the simulation.

---

lie in the thousands, and the minimum infective dose for *E. coli* in the tens to millions[1]. Accordingly, making the assumption that strains disappear once they reach 3% of the frequency of the typical germ size—where the typical germ size may well be larger than the *minimum* infective dose—is likely to be a safe estimate.

### 1.4 Illustration of the correspondence between individual-based and ODE implementations

To demonstrate how the two alternative model implementations produce very similar results, we show overlapping results from the individual-based and ODE model implementations in **Fig. S2**.

## Supplementary Note 2. Dual carriage promotes coexistence

*Overview — In this supplementary note, we identify model parameters that impact upon the rate of dual carriage (i.e. carriage of both sensitive and resistant strains) and, in doing so, modulate the extent of coexistence between sensitive and resistant strains. We conclude that it is the rate of dual carriage per se that determines the extent of coexistence, and that strain knockout within hosts (which occurs in the knockout model) inhibits coexistence at the population level.*

*Efficiency of co-colonisation and knockout* — In the main text, we interpret the parameter $k$ as the relative efficiency of co-colonisation compared to primary colonisation. We show that as $k$ increases, coexistence increases across both the knockout and mixed-carriage models. We describe $k$ as the conditional probability of successful co-colonisation given the transmission of a germ to a host who is already colonised. However, $k$ can also be interpreted more generally as a multiplier on the base rate of colonisation, such that values of $k > 1$ represent a scenario in which carriers are more likely than non-carriers to be newly (co-)colonised.

The parameter $k$ can potentially summarize multiple phenomena. For example, resident strains may interfere with an incoming strain's ability to establish itself within the host through competition or because they have activated host immunity, either of which could inhibit co-colonisation, effectively reducing $k$. Alternatively, the resident strain may induce inflammation of host tissues, which could promote acquisition of further strains, effectively increasing $k$. It is also possible to interpret $k$ as capturing increased contact among carriers compared to non-carriers, and hence values of $k > 1$ could capture higher than expected transmission among individuals who are prone to carriage, standing in for a "population-structuring" effect whereby individuals who are more prone to acquiring carriage tend to associate preferentially with each other. There is often good evidence for this phenomenon—for example, children are more susceptible to colonisation by *S. pneumoniae* than adults[13], and children are also more likely to make physical contact with other children than with adults.

Since the amount of co-colonisation increases with $k$, higher values of $k$ might, in theory, allow the knockout model to account for more coexistence. However, as we show in **Fig. S3a**, increasing $k$ in the knockout model also increases the rate of knockout, such that overall, increases to $k$ even above $k = 1$ do not substantially increase coexistence for the knockout model. Increasing $k$ does, however, have a comparatively greater effect on the extent of coexistence in the mixed-carriage models (**Fig. S3b&c**).

*Germ size* — Another difference between the knockout model and the mixed-carriage model is that the knockout model assumes that a successfully co-colonising strain reaches a within-host frequency of $1/2$, while the mixed-carriage model assumes that co-colonising strains are initially present at a within-host frequency of $\iota/(1+\iota)$. This might potentially impact upon the relative extent of coexistence shown by each model. However, as we show in **Fig. S3d**, when we set $\iota = 1$ in the mixed-carriage model (making the within-host frequency of newly co-colonised strains $1/2$, the same as in the knockout model) there is almost no impact upon the potential for coexistence.

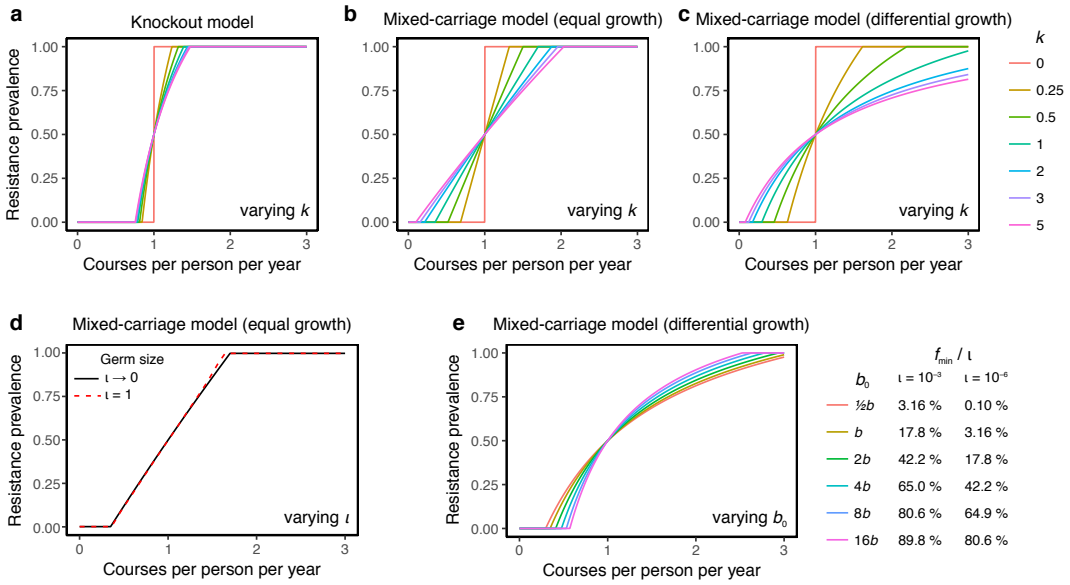



**Fig. S3 | Factors inhibiting or promoting coexistence. (a–c)** Increasing $k$ above 1 promotes coexistence, but the effect is considerably smaller in the knockout model **(a)** than in the mixed-carriage models **(b, c)**. Parameters are otherwise the same as in Fig. 3c, f, i in the main text. **(d)** In the mixed-carriage model with equal within-host strain growth, increasing the germ size from negligably small ($\iota \to 0$, solid black line, ODE implementation) to large ($\iota = 1$, dashed red line, individual-based implementation) has very little impact upon coexistence. Other parameters are $\beta = 5, k = 1, u = 1$, and $c = 0.124$ (black line) versus $c = 0.0456$ (red dashed line), chosen such that resistance prevalence passes through 0.5 at $\tau = 1$ y$^{-1}$. **(e)** Increasing $b_0$ from its normal value of $b/2$ decreases the potential for coexistence, but $b_0$ must be increased substantially to have a major impact upon how much coexistence is exhibited by the model. On the right side of the figure, the equivalent $f_{\min}$, expressed as a relative percentage of the germ size $\iota$, is shown for two different values of the germ size. For example, if we assume $b_0 = b/2$, then for $\iota = 10^{-3}$ we are assuming that a strain disappears once it decreases to ~3% of its germ size, and for $\iota = 10^{-6}$ we are assuming that a strain disappears once it decreases to ~0.1% of its germ size. Other parameters as in **(c)**, with $b$ chosen so that resistance prevalence passes through 0.5 at $\tau = 1$ y$^{-1}$.

---

*Within-host competitive exclusion* — Finally, we test the extent to which within-host competitive exclusion of resistant strains by sensitive strains owing to within-host growth of sensitive cells impacts upon coexistence. In **Fig. S3e**, we show that increasing $b_0$ reduces the amount of coexistence exhibited by the model. However, $b_0$ must be increased substantially in order to appreciably reduce coexistence.

All in all, these findings suggest (i) that it is the strain-knockout property of the knockout model that inhibits coexistence in particular; (ii) that increased co-colonisation promotes coexistence in all models, so factors that increase co-colonisation (such as greater $k$) promote coexistence while factors that decrease co-colonisation (such as greater $b_0$) inhibit coexistence; and (iii) that increasing $k$ has a comparatively smaller impact upon coexistence in the knockout model compared to the mixed-carriage model because $k$ not only leads to the creation of dual carriers, but simultaneously depletes the population of dual carriers through strain knockout.

In support of point (iii) above, note that when we re-fit all models to empirical data allowing $k$ to exceed 1 (specifically, adopting a uniform prior over $0 \leq k \leq 5$ instead of

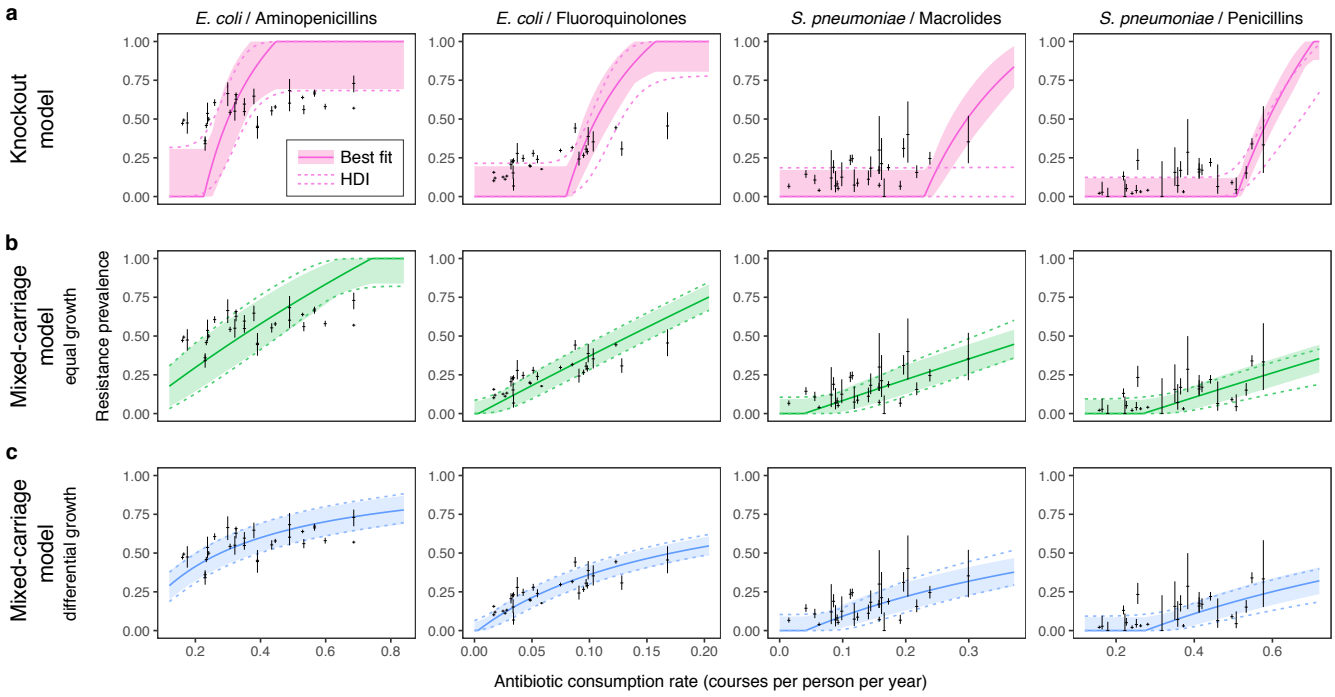



**Fig. S4 | Model fits when we allow *k* to exceed 1 ($0 \leq k \leq 5$).** Fits are not markedly different than when $0 \leq k \leq 1$ (see Fig. 4, main text). Solid lines and ribbons show the single best-fit run for each model (solid lines) and the 67% highest density interval incorporating between-country random effects (ribbon). Regions bounded by dashed lines show the 67% HDI across the estimated posterior, again incorporating between-country random effects.

---

$0 \leq k \leq 1$), the fit of the knockout model is not substantially improved (**Fig. S4**). Further details of this model fitting scenario are given in **Supplementary Note 4**.

**Fig. S5** illustrates more directly the relationship between the frequency of dual carriage and the amount of coexistence. We constructed this figure by fitting the mixed-carriage model with differential within-host growth for a fixed $\beta$ and $u$ ($\beta = 2, u = 1$ for *S. pneumoniae* and $\beta = 2, u = 0.25$ for *E. coli*), with a uniform prior on $k$ from 0 to 25, and multiplying the likelihood by a penalty $\mathcal{P} = \text{Beta}\big(d(\theta)\big|\alpha = 1000d^*, \beta = 1000(1 - d^*)\big)$ where $\text{Beta}(x|\alpha, \beta)$ is the beta distribution PDF, $d(\theta)$ is the fraction of carriers carrying both sensitive and resistant strains (*i.e.* $\frac{1-X-S-R}{1-X}$), and $d^*$ is a "target" fraction of dual carriers. Effectively, this forces the fraction of dual carriers to be close to $d^*$, and illustrates our assertion that it is the fraction of dual carriers—rather than other parameters such as the transmission rate or rate of co-colonisation *per se*—which determines the extent of coexistence in the model. We show $d^* = 0.6, 0.4, 0.2$ for *E. coli* and $d^* = 0.3, 0.2, 0.1$ for *S. pneumoniae*.

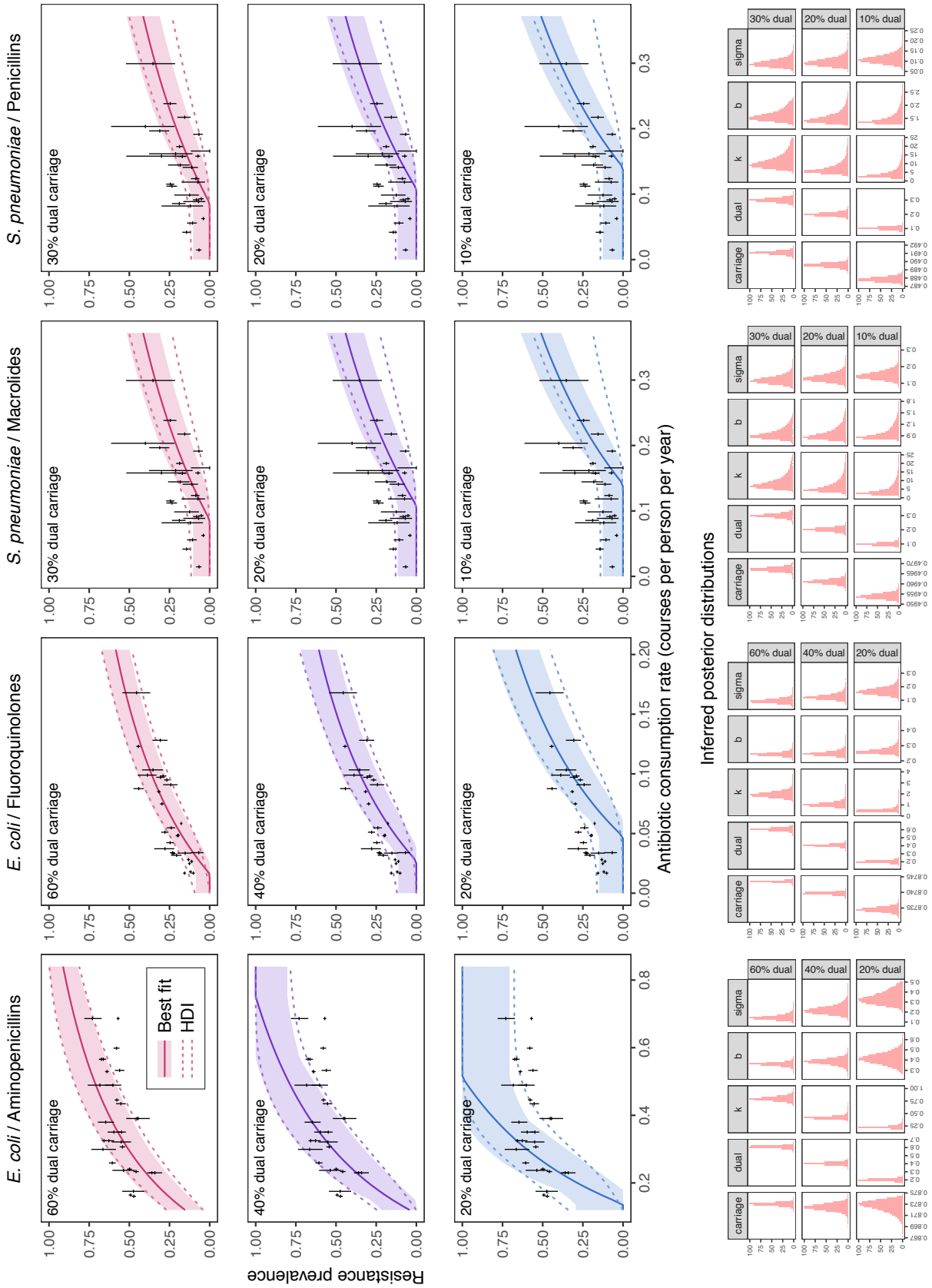



**Fig. S5 | Model fits when dual carriage is constrained.** Inferred posterior distributions are shown to illustrate constraints on dual carriage (panels marked "dual") and how other parameters change due to constraints on dual carriage: "k" and "sigma" show larger relative shifts, particularly for *E. coli*, while "carriage" and "b" show smaller shifts. See text for details.

## Supplementary Note 3. The mixed-carriage model is structurally neutral

*Overview — In the main text, we introduce the mixed-carriage model and claim that it is structurally-neutral. We support this assertion in this supplementary note.*

*Section 3.1 gives an intuitive argument that the mixed-carriage model is structurally neutral, then further argues that the mixed-carriage model meets the two key criteria for structural neutrality established by Lipsitch et al.[3]: "ecological neutrality" and "population-genetic neutrality". It also shows graphically that the mixed-carriage model, when analysing the dynamics of two equivalent strains, keeps relative strain frequencies unchanged over time irrespective of initial conditions.*

*Section 3.2 discusses how structurally-neutral models may or may not exhibit within-host neutrality.*

### 3.1 Structural neutrality of the mixed-carriage model

Intuitively, a structurally-neutral model is one in which, when multiple equivalent strains are being analysed, all model dynamics are essentially unbiased with respect to the identities of the strains being analysed, such that the model dynamics are governed entirely by unbiased random sampling of individual pathogens (*i.e.*, by drift).

Suppose that two equivalent strains, A and B, were analysed with the knockout model. Since these strains are equivalent, assume without loss of generality that they are both unaffected by antibiotic treatment. Alternatively, they could both be affected by antibiotic treatment, which then becomes indistinguishable from an inflated rate of natural clearance, $u' = u + \tau$. In the knockout model, when a non-carrier is colonised, the probability that it becomes colonised with strain A is equal to the relative frequency of strain A in the population, while the probability that it instead becomes colonised with strain B is equal to the relative frequency of strain B in the population. Therefore, colonisation is neutral with respect to strain identities. When a carrier is co-colonised, the contents of one of its two subcompartments is replaced with either strain A or strain B, again proportionally to the relative frequency of that strain in the population. Accordingly, co-colonisation is also neutral with respect to strain identities. Finally, carriers undergo natural clearance irrespective of the actual strains they are carrying in either subcompartment, so clearance is also neutral with respect to strain identities. In summary, when analysing equivalent strains, the knockout model's dynamics are governed entirely by drift, which shows that the knockout model is structurally neutral.

The mixed-carriage model is structurally neutral for similar reasons. Suppose we were to use the mixed-carriage model to analyse equivalent strains. This requires that we assume no differential within-host growth and that $f_{\min} = Y_{\min} = 0$ to prevent strain identities from having any impact upon model dynamics. Then, colonisation is neutral with respect to strain identities because when a non-carrier is colonised, the strain they are colonised with is chosen with probability equal to its population-level frequency. Co-colonisation is also neutral because it replaces a fraction $\iota/(1 + \iota)$ of cells in a carrier with cells of a random strain, also chosen with probability equal to that strain's

population-level frequency. And clearance is neutral because hosts experience clearance events independently of the mix of strains they are carrying. For that reason, when analysing equivalent strains, the mixed-carriage model is governed entirely by drift and is therefore structurally neutral.

This argument can be made more rigorous. We argue below that the mixed-carriage model meets the two criteria for structural neutrality proposed by Lipsitch *et al.*[3]—"ecological neutrality" and "population-genetic neutrality"—and hence is structurally neutral.

3.1.1 Ecological neutrality

In order for a model to be ecologically neutral for identical strains, it must be possible to rewrite the model in terms of "ecological state variables"—namely, the number of uninfected hosts and the number of hosts that have been colonised 0, 1, 2, *etc.*, times—in a way which is independent of identities of any particular strains involved[3]. To meet the assumption of indistinguishable strains, we set $c = 0$, $\tau = 0$ and $w_s = 1$, and in order to prevent neutral labels from having an impact upon strain dynamics, we assume that $f_{\min} = Y_{\min} = 0$. Now note that we can rewrite the mixed-carriage model as a series of transitions between host states $N_0, N_1, N_2, \ldots$ defined by the subscript $M$, the multiplicity of infection (*i.e.*, the total number of colonisations experienced by a specific host since their last episode of natural clearance):

$$N_0 \xrightarrow{\lambda_{\text{tot}}} N_1 \quad (\textit{colonisation})$$
$$N_M \xrightarrow{k\lambda_{\text{tot}}} N_{M+1} \quad \text{for all } M > 0 \quad (\textit{co-colonisation})$$
$$N_M \xrightarrow{u} N_0 \quad \text{for all } M > 0 \quad (\textit{clearance}),$$

where $\lambda_{\text{tot}} = \beta(N_1 + N_2 + \cdots + N_\infty)$ is the total force of infection in the population. This is enough to fully specify the model if we are indifferent to the identities of the indistinguishable strains that are circulating.

Note that the within-host frequency $f_m$ attributable to the $m$th colonising strain in a host that has been colonised $M$ times is

$$f_1 = \frac{1}{(1+\iota)^{M-1}},$$
$$f_m = \frac{\iota}{(1+\iota)^{M-m+1}} \quad \text{for all } m \geq 2.$$

3.1.2 Population-genetic neutrality

In order for a model with two strains to meet the criterion of population-genetic neutrality, the expected frequency of either strain should not change over time if the two strains are identical apart from a biologically-meaningless label[3]. For the mixed-carriage model, this means that both strains will have equal within-host fitness (meaning that within-host growth can be neglected; see above) and either that both

strains are resistant (in which case treatment has no effect, so we can assume $\tau = 0$) or both strains are sensitive (in which case treatment and natural clearance can be treated together as clearance at rate $u' = u + \tau$). We will also assume $f_{\min} = 0$ and $Y_{\min} = 0$.

Since we are free to ignore treatment and within-host growth, this means that the only changes to a given population will occur through a random sequence of transmission and clearance events. Let us refer to the two strains of the model as strain A and strain B. Each random clearance or transmission event will cause a small perturbation to the frequency of strain A, and an equal and opposite perturbation to the frequency of strain B. Our aim here is twofold. First, we will show that the expected value of these perturbations to the frequency of strain A is zero for each type of event regardless of the state of the population. In doing so, we will show that the mixed-carriage model does not favour either strain A or B arbitrarily. Second, we will show that the magnitude (*i.e.* absolute value) of any such perturbation goes to zero as the number of carriers goes to infinity. This shows that as the total population size goes to infinity, the combined effect of all transmission and clearance events in a fixed time period goes to zero, and hence the mixed-carriage model satisfies population-genetic neutrality, suggesting that any stochastic fluctuations for a finite population are attributable to drift.

Suppose that there are *N* hosts in total, *K* of which are carriers (the remaining *N* – *K* are non-carriers). Of the *K* carriers, the *i*th carrier's carriage of strain A is $x_i$ and their carriage of strain B is $1 - x_i$. The overall frequency of strain A in the population is $X = \frac{1}{K}\sum_{i=1}^{K} x_i$, while the total carriage of strain A is $KX = \sum_{i=1}^{K} x_i$; note that *X* is also the expected value of $x_i$ for a random carrier, since $\mathrm{E}(x_i) = \sum_{i=1}^{K} \frac{1}{K} x_i = \frac{1}{K}\sum_{i=1}^{K} x_i = X$. If the frequency of strain A before some event is *X*, and the frequency of strain A following the event is $X'$, our aim is (1) to show that $\mathrm{E}(X') = X$ for both clearance and transmission events and (2) that the magnitude of any of these perturbations is inversely proportional to the number of carriers, *i.e.* $|X' - X| \propto \frac{1}{K}$.

*Clearance* — When clearance occurs, a random carrier *j* has their carriage eliminated, which means the number of carriers, *K*, decreases by 1 and the total population carriage of strain A, *KX*, decreases by $x_j$. Therefore, the expected frequency of strain A following a clearance event is

$$\begin{aligned}\mathrm{E}(X') &= \mathrm{E}\left(\frac{KX - x_j}{K-1}\right)\\ &= \frac{KX - \mathrm{E}(x_j)}{K-1}\\ &= \frac{KX - X}{K-1}\\ &= \frac{(K-1)X}{K-1} = X,\end{aligned}$$

*i.e.* clearance leaves the expected frequency of strain A unchanged. Note that any one clearance event changes the frequency of strain A by

$$\frac{KX - x_j}{K-1} - X = \frac{X - x_j}{K-1},$$

which goes to zero as $K \to \infty$.

*Transmission* — There are two types of transmission events: transmission events which result in the colonisation of uncolonised hosts, and transmission events which result in the colonisation of already-colonised hosts. For the first type of transmission event, the probability that strain A is being transmitted is $X$ and the probability that strain B is being transmitted is $1 - X$. The outcome is that an extra carrier is added, such that the total number of carriers becomes $K + 1$, and the new carrier is a strain-A carrier with probability $X$ (increasing total carriage of strain A by 1) and a strain-B carrier with probability $1 - X$ (keeping total carriage of strain A the same). Therefore the expected frequency of strain A following a transmission event to an uncolonised host is

$$\begin{aligned} \mathrm{E}(X') &= X\left(\frac{KX+1}{K+1}\right) + (1-X)\left(\frac{KX}{K+1}\right) \\ &= \frac{KX^2 + X + KX - KX^2}{K+1} \\ &= \frac{X(K+1)}{K+1} = X\,, \end{aligned}$$

*i.e.* transmission to an uncolonised host leaves the expected frequency of strain A unchanged. Note that any single transmission to an uncolonised host changes the frequency of strain A by

$$\frac{KX+1}{K+1} - X = \frac{1-X}{K+1}$$

if strain A is being transmitted and

$$\frac{KX}{K+1} - X = -\frac{X}{K+1}$$

if strain B is being transmitted, which both go to zero as $K \to \infty$.

Finally, if a carrier $j$ experiences a transmission event, their carriage of strain A will change from $x_j$ to $\frac{x_j+\iota}{1+\iota}$ if strain A is being transmitted and to $\frac{x_j}{1+\iota}$ if strain B is being transmitted, where $\iota$ is the germ size. Equivalently, carrier $j$'s carriage of strain A changes by $\frac{x_j+\iota}{1+\iota} - x_j = \frac{\iota}{1+\iota}(1 - x_j)$ with probability $X$, and changes by $\frac{x_j}{1+\iota} - x_j = -\frac{\iota}{1+\iota}x_j$ with probability $1 - X$. If a single carrier's strain-A carriage changes by some amount $y$,

then the population frequency of A changes by $y/K$; overall, the expected population-level frequency of strain A following a transmission event to a colonised host is

$$\begin{aligned}
\mathrm{E}(X') &= \mathrm{E}\left( X\left(X + \frac{\iota}{1+\iota}\frac{1-x_j}{K}\right) + (1-X)\left(X - \frac{\iota}{1+\iota}\frac{x_j}{K}\right)\right) \\
&= X\left(X + \frac{\iota}{1+\iota}\frac{1-\mathrm{E}(x_j)}{K}\right) + (1-X)\left(X - \frac{\iota}{1+\iota}\frac{\mathrm{E}(x_j)}{K}\right) \\
&= X\left(X + \frac{\iota}{1+\iota}\frac{1-X}{K}\right) + (1-X)\left(X - \frac{\iota}{1+\iota}\frac{\mathrm{X}}{K}\right) \\
&= X^2 + \frac{\iota}{1+\iota}\frac{X(1-X)}{K} + X - X^2 - \frac{\iota}{1+\iota}\frac{X(1-X)}{K} = X\,,
\end{aligned}$$

*i.e.* transmission to a colonised host also leaves the expected frequency of strain A unchanged. Note that, as stated above, any single transmission to a colonised host changes the frequency of strain A by $\frac{\iota}{1+\iota}\frac{1-x_j}{K}$ if strain A is being transmitted and $-\frac{\iota}{1+\iota}\frac{x_j}{K}$ if strain B is being transmitted, and both of these go to zero as $K \to \infty$.

Since the expected value of any perturbation to strain frequencies is zero, and the magnitude of any one perturbation to strain frequencies goes to zero as the number of carriers goes to infinity, the mixed-carriage model exhibits population-genetic neutrality when the population size is infinite.

3.1.3 Dynamics of the mixed-carriage model

We can also informally illustrate the population-genetic neutrality of the mixed-carriage model graphically — note that, when strains are identical (*i.e.* $\tau = 0, c = 0$), the ODE-based model retains the strain frequencies it begins with (**Fig. S6**, top row).

### 3.2 Within-host neutrality

As we argue in the main text, the knockout model meets the criteria for structural neutrality proposed by Lipsitch *et al.*[3], but violates the spirit of structural neutrality by assuming that all cells from one of two "subcompartments" are eliminated from carriage during knockout. Whether this, in fact, is compatible with the idea of structural neutrality depends upon the interpretation of "SR" (dual-strain) carriers in the model.

One possibility is that hosts really are subdivided into two physically distinct subcompartments which can, for whatever reason, only be occupied by one strain at a time. In this case it would make sense for knockout to eliminate all of the cells in one of the two host subcompartments, but there are clear difficulties interpreting what these subcompartments might physically correspond to, and moreover there is ample evidence that individuals can carry two strains or more in a single physical niche[4–7]. Another interpretation is that hosts are only capable of carrying up to two bacterial cells at once. In this case, it is possible that an invading cell might only replace one of the two cells. Although this is an obviously unrealistic scenario, it illustrates how the neutrality of a model can partly depend upon the interpretation of host states. Finally, a third

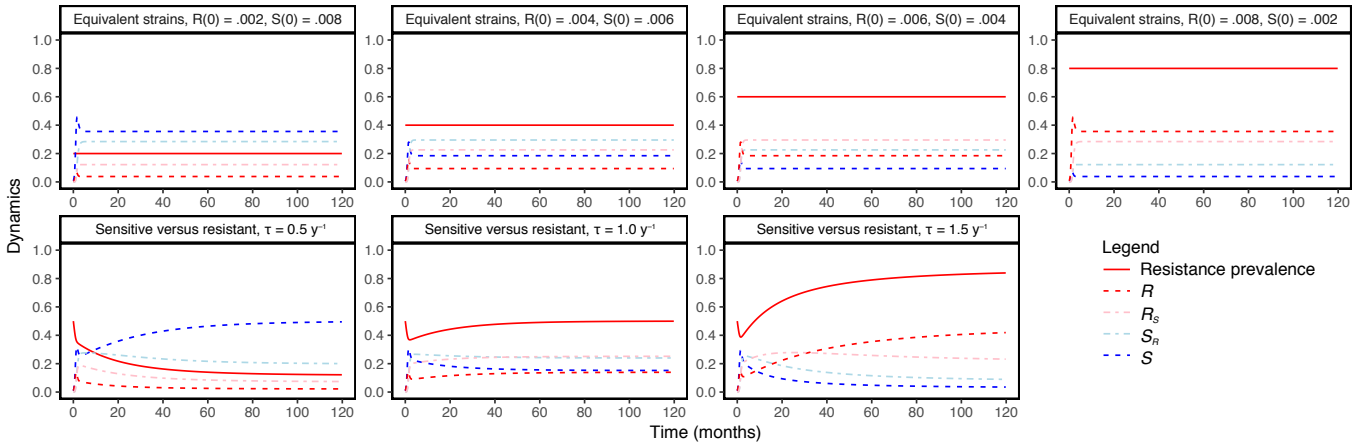



**Fig. S6 | Dynamics of the ODE implementation of the mixed-carriage model.** *Top row*: Model dynamics for the mixed-carriage ODE model where the $S$ strain and $R$ strain are equivalent. That is, we have $\tau = 0,\ c = 0$, and $b = 0$. Other parameters are $\beta = 5,\ u = 1$, and $k = 1$. Each panel of the top row shows the same parameters but different initial frequencies of S and R carriers. Note that the relative prevalence of the $R$ strain—*i.e.* $(R + R_S)/(R + R_S + S_R + S)$—stays constant over time, remaining equal to initial prevalence at $t = 0$ as shown in the figure headings. *Bottom row:* example dynamics for non-equivalent strains are illustrated. These correspond to Fig. 3f of the main text, *i.e.* $\beta = 5,\ c = 0.124,\ b = 0,\ k = 1,\ u = 1$, and $\tau$ as given in the figure heading.

---

possibility is that hosts comprise a single niche, and SR carriers represent hosts in which the niche carries half resistant cells and half sensitive cells. This interpretation is incompatible with structural neutrality, because even when S cells and R cells only differ by a biologically-meaningless marker, they are eliminated en bloc during knockout.

In summary, we argue that the knockout model cannot simultaneously be used to model transmission dynamics among hosts capable of carrying a large number of diverse pathogens in the same niche, while also adhering to the motivating concept of structural neutrality which dictates that model dynamics should not be influenced by a neutral label applied to some subset of pathogens. We suggest that models incorporating within-host dynamics should endeavour to treat individual pathogens (whether microbes, viruses, or macroparasites) neutrally, rather than only treating strains neutrally.

## Supplementary Note 4. Model fitting details

*Overview — In this supplementary note, further details are given of the model fitting procedure used in the main text.*

### 4.1 Prior distributions for model fitting

**Table S1** summarises prior distributions used in model fitting. Note that for *S. pneumoniae*, we assume an average duration of carriage of 1 month for consistency with previous studies[2,8], while for *E. coli*, we assume that the average duration of carriage is 59 to 98 days[9,10], and accordingly set a uniform prior for $u$ over the range 0.3–0.5 months$^{-1}$. The transmission rate $\beta$ is indirectly constrained by a likelihood penalty on prevalence of carriage, so we set a uniform prior for $\beta$ wide enough to overlap the full range of permissible carriage prevalence, $Y$, given the range of clearance rates $u$ and treatment rates $\tau$ (*i.e.* $Y = 1 - (u + \tau)/\beta$ for the sensitive strain alone, and $Y = 1 - u/\big(\beta(1-c)\big)$ for the resistant strain alone).

| | *E.coli* / Aminopenicillins: 5 parameters | *E. coli* / Fluoroquinolones: 5 parameters | *S. pneumoniae* / Macrolides: 4 parameters | *S. pneumoniae* / Penicillins: 4 parameters |
|---|---|---|---|---|
| | Fitted parameters | | | |
| $\beta$ (transmission rate) | 0.75 – 10 mo$^{-1}$ | 0.75 – 10 mo$^{-1}$ | 1 – 6 mo$^{-1}$ | 1 – 6 mo$^{-1}$ |
| $c$ (transmission cost of resistance: knockout & equal-growth mixed-carriage models only) | 0 – 1 | 0 – 1 | 0 – 1 | 0 – 1 |
| $b$ (within-host growth benefit of sensitivity: differential-growth mixed-carriage model only) | 0 – 10 | 0 – 10 | 0 – 10 | 0 – 10 |
| $k$ (relative efficiency of co-colonisation) | 0 – 1 | 0 – 1 | 0 – 1 | 0 – 1 |
| $u$ (natural clearance rate) | 0.3 – 0.5 mo$^{-1}$ | 0.3 – 0.5 mo$^{-1}$ | fixed (1 mo$^{-1}$) | fixed (1 mo$^{-1}$) |
| $\sigma$ (additional between-country variability in resistance prevalence) | 0 – 1 | 0 – 1 | 0 – 1 | 0 – 1 |
| | Likelihood components | | | |
| $Y$ (prevalence of carriage) | 0.499 – 0.942 | 0.499 – 0.942 | 0.3 – 0.8 | 0.3 – 0.8 |

**Table S1 | Priors used in model fitting.** All priors are uniform over the ranges specified.

In the Methods, we detail how the likelihood function used in model fitting constrains model output such that all countries must exhibit a prevalence of carriage $Y$ such that $Y^{(0)} \leq Y \leq Y^{(1)}$. For *S. pneumoniae*, we follow Colijn *et al.*[2] in assuming $0.3 \leq Y \leq 0.8$ in children. Carriage of *E. coli* is essentially universal, but because we are interested in strains that can potentially cause invasive disease, we restrict our attention to extraintestinal pathogenic *E. coli* (ExPEC), a subset of *E. coli* that is responsible for most invasive infections. We assume that the carriage of ExPEC is in the range $0.499 \leq Y \leq 0.942$, which corresponds to 95% confidence intervals around the observed prevalence of carriage of ExPEC in a study by Martinez-Medina *et al.*[6] (which found that ExPEC was carried by 9 out of 12 healthy subjects).

### 4.2 Details of MCMC

We use the differential evolution MCMC algorithm[11], running $10n$ chains, where $n$ is the number of free parameters in the model, *i.e.* $10n = 40$ for *S. pneumoniae* (for which carriage duration is fixed at 1 month) and $10n = 50$ for *E. coli* (for which carriage duration is not fixed). The burn-in period lasts 1,000 iterations, after which 100,000 samples from the posterior are taken across all chains. MCMC convergence and effective sample sizes, calculated using the R package *coda*[12], are in **Appendix S1**.

### 4.3 Posterior distributions from model fitting

Posterior distributions from model fitting are shown in **Figs. S7** and **S8**. Pairwise joint distributions for the main analysis are shown in **Appendix S2**.

### 4.4 Model fitting assessment

We use AIC in the main text to formally assess model fit. Deviance, defined as $-2\mathcal{L}$, where $\mathcal{L}$ is the likelihood, is an alternative way of assessing model fit which gives a distribution rather than a single value. We provide 95% HDIs for the deviance of each model fit in **Appendix S3**. Note that the mixed-carriage model with and without within-host growth are more comparable for *S. pneumoniae* than when $0 \leq k \leq 1$.

---

**Fig. S7 (next page) | Posterior distributions for model fitting when $0 \leq k \leq 1$ (i.e., from the main text).** Here, *carriage* gives the overall prevalence of carriage in the population; *dual* gives the fraction of carriers who carry both sensitive and resistant strains; *beta* gives the transmission rate; *c* gives the transmission cost of resistance; *u* gives the clearance rate (if the clearance rate is subject to fitting; for *S. pneumoniae*, $u = 1$); *k* gives the relative efficiency of co-colonisation; *b* gives the within-host growth rate of the sensitive strain; and *sigma* gives the standard deviation of unexplained between-country variation in resistance prevalence. Knockout: knockout model; Mixed / Equal: mixed-carriage model with equal within-host growth; Mixed / Diff: mixed-carriage model with differential within-host growth. Note that each histogram has been scaled to the full height of the panel.

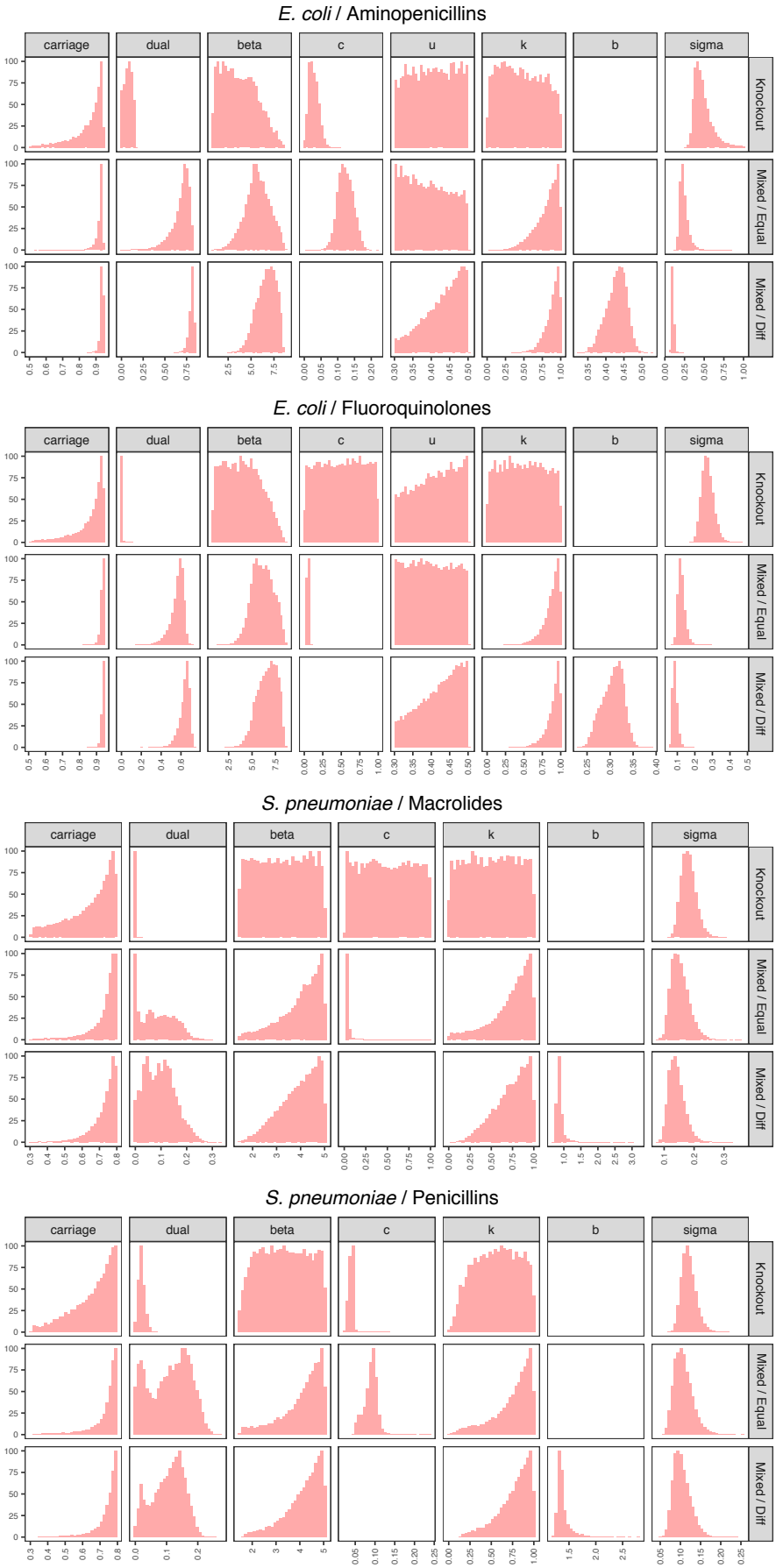



**Fig. S7 |** See previous page for caption.

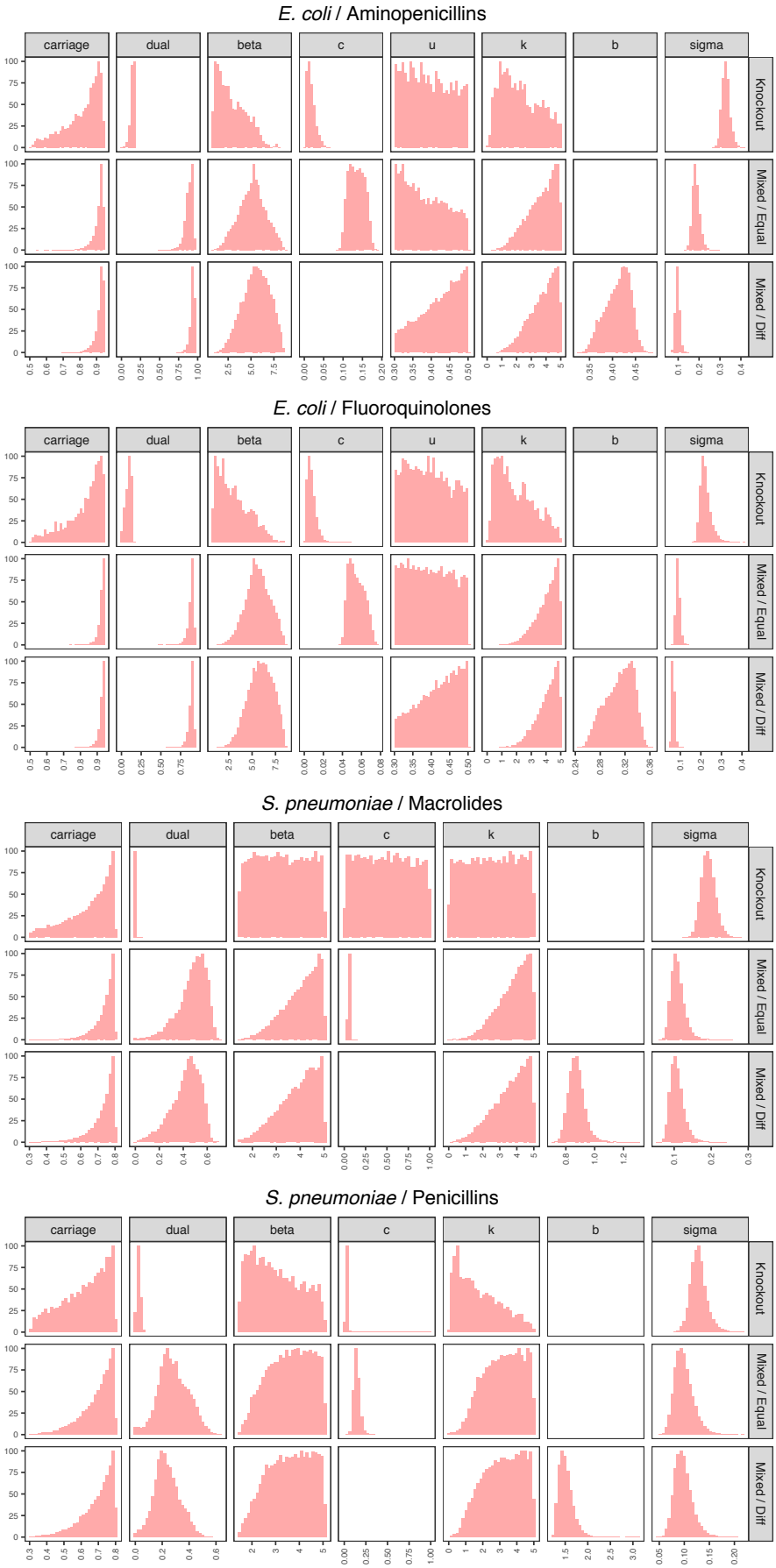



**Fig. S8 | Posteriors for model fitting when $0 \leq k \leq 5$.** See Fig. S7 for details.

## Supplementary Note 5. The mixed-carriage model with multiple serotypes and host immunity

*Overview — In the main text, we present results from an extended mixed-carriage model that allows us to analyse dynamics of multiple serotypes (i.e., more than two strains at a time) and host adaptive immunity. In this supplementary note, we describe how the model is implemented and show results for our analysis of resistance evolution among pneumococcal serotypes in the absence of adaptive immunity.*

Here, we provide details of the "extended" mixed-carriage model that can accommodate any number of strains. In this individual-based model implementation, there are $N$ hosts, $L$ serotypes, and $M = 2L$ strains. The frequency of host $i$'s carriage of strain $j$ is $f_{i,j}$, and a host's total carriage is $F_i = \sum_j f_{i,j}$. We assume that strains 1 and 2 are of serotype 1, strains 3 and 4 are of serotype 2, strains 5 and 6 are of serotype 3, and so on, and that odd-numbered strains are sensitive while even-numbered strains are resistant. Two different kinds of process act upon hosts: "updates" to within-host growth occur at discrete time intervals of $\Delta t = 0.001$, while transmission, clearance, and treatment events are Poisson processes that occur at random times between updates.

During updating, any strains which have a frequency of less than $f_{\min}$ are cleared; then each strain in each carrier grows by a factor $\omega_j = w_j^{\Delta t}$, where $w_j$ is strain $j$'s per-unit-time within-host growth rate; then each carrier's total carriage is normalised so that $F_i = 1$. That is,

$$\left(f_{i,1}, f_{i,2}, \ldots, f_{i,M}\right) \to \left(\frac{\omega_1 q\left(f_{i,1}\right)}{\sum_j \omega_j q\left(f_{i,j}\right)}, \frac{\omega_2 q\left(f_{i,2}\right)}{\sum_j \omega_j q\left(f_{i,j}\right)}, \ldots, \frac{\omega_M q\left(f_{i,M}\right)}{\sum_j \omega_j q\left(f_{i,j}\right)}\right),$$

where

$$q(a) = \begin{cases} a & \text{if } a \geq f_{\min} \\ 0 & \text{if } a < f_{\min} \end{cases}.$$

Note that if all carried strains have a frequency of less than $f_{\min}$, then the right-hand side of the transition notated above evaluates to $\left(\frac{0}{0}, \frac{0}{0}, \ldots, \frac{0}{0}\right)$. In this case, we set a host's state to $(0,0, \ldots, 0)$.

The force of infection for each strain $j$ is $\lambda_j = \beta_j \max\left(Y_{\min}, \sum_i f_{i,j}\right)/N$ (we can set $Y_{\min} = 1$ to effectively assume there is always at least one carrier of each strain in order to avoid stochastic elimination of strains[13], or set $Y_{\min} = 0$ to not do this). Here, $\beta_j$ is the transmission rate for strain $j$, including any transmission-rate penalty for resistance—that is, for a two-strain model with a sensitive and a resistant strain, we could write $\beta_1 = \beta, \beta_2 = \beta(1 - c)$. Events comprise transmission events, clearance events, and treatment events. Specifically: transmission events for each strain $j$ occur at rate $\kappa_i \lambda_j$ to each host, where $\kappa_i = 1$ if $F_i = 0$ and $\kappa_i = k$ if $F_i > 0$; clearance events for each serotype $\ell$ occur at rate $u_\ell$ to each host, where $u_\ell$ is the clearance rate for serotype $\ell$; and

treatment events occur at rate $\tau$ to each host, where $\tau$ is the antibiotic treatment rate. Events have the following effect on hosts:

$$\begin{aligned}
&(f_{i,j}) \xrightarrow{\kappa_i \lambda_j} (f_{i,j} + \iota) \quad (transmission)\\
&(f_{i,2\ell-1}, f_{i,2\ell}) \xrightarrow{u_\ell} (0,0) \quad (clearance)\\
&(f_{i,1}, f_{i,3}, f_{i,5}, \dots, f_{i,M-1}) \xrightarrow{\tau} (0,0,0, \dots ,0) \quad (treatment),
\end{aligned}$$

where each of the above transitions is immediately followed by the transition

$$(f_{i,1}, f_{i,2}, \dots, f_{i,M}) \to \left(\frac{f_{i,1}}{F_i}, \frac{f_{i,2}}{F_i}, \dots, \frac{f_{i,M}}{F_i}\right)$$

if $F_i > 0$, to re-enforce carrying capacity. Above, we only notate the components of host carriage that may change for each event; that is, transmission of strain *j* only affects $f_{i,j}$ initially (*i.e.*, prior to enforcement of carrying capacity); clearance of serotype $\ell$ only affects $f_{i,2\ell-1}$ and $f_{i,2\ell}$ initially; and treatment only affects the sensitive strains of each serotype (*i.e.*, odd-numbered strains) initially.

When serotype-specific adaptive immunity is introduced, we introduce birth events, which occur at rate $\alpha$ (*i.e.*, the birth rate) for each host, and we also keep track of immunities $m_{i,\ell}$, where $m_{i,\ell} = 1$ if host *i* is immune to serotype $\ell$ and $m_{i,\ell} = 0$ if host *i* is not immune to serotype $\ell$. Immunity to a serotype is gained when hosts naturally clear that serotype, and immunity confers total protection against future colonisation by that serotype. Birth represents the entry of new, immunologically-naïve and uncolonised hosts into the set of potentially-susceptible hosts and the simultaneous departure of older hosts. Events are now

$$\begin{aligned}
&(f_{i,j}) \xrightarrow{(1-m_{i,\lceil j/2 \rceil})\kappa_i \lambda_j} (f_{i,j} + \iota) \quad (transmission)\\
&(f_{i,2\ell-1}, f_{i,2\ell}); (m_{i,\ell}) \xrightarrow{u_\ell} (0,0); (1) \quad (clearance)\\
&(f_{i,1}, f_{i,3}, f_{i,5}, \dots, f_{i,M-1}) \xrightarrow{\tau} (0,0,0, \dots ,0) \quad (treatment)\\
&(f_{i,1}, f_{i,2}, \dots, f_{i,M}); (m_{i,1}, m_{i,2}, \dots, m_{i,\ell}) \xrightarrow{\omega} (0,0, \dots ,0); (0,0, \dots ,0) \quad (birth),
\end{aligned}$$

where $\lceil \cdot \rceil$ is the ceiling function, and each of the above transitions is immediately followed by the transition

$$(f_{i,1}, f_{i,2}, \dots, f_{i,M}) \to \left(\frac{f_{i,1}}{F_i}, \frac{f_{i,2}}{F_i}, \dots, \frac{f_{i,M}}{F_i}\right)$$

if $F_i > 0$, to re-enforce carrying capacity. That is, transmission of serotype $\ell$ is blocked if the host is immune to that serotype; clearance of serotype $\ell$ by host *i* makes host *i* immune to serotype $\ell$ does not affect immunity to other serotypes; and birth replaces host *i* with a new host that carries no strains and is immune to no serotypes. At the start of the simulation, all $m_{i,\ell} = 0$.

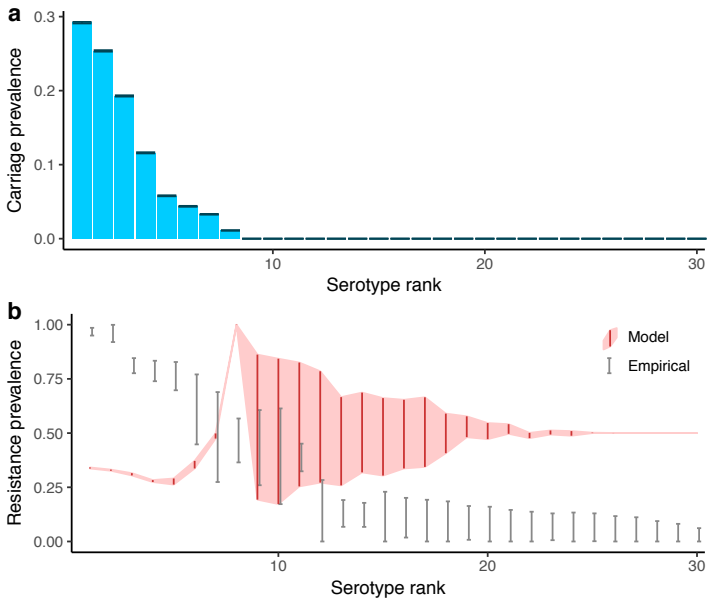



**Fig. S11 | Immunity is needed to maintain significant serotype diversity.** As serotypes vary greatly in duration of carriage, the mechanism of serotype-specific clearance alone is not able to reproduce observed patterns of pneumococcal carriage or resistance prevalence, as all but the 8 serotypes with the highest duration of carriage are eliminated. Resistance prevalence is close to 50% for eliminated serotypes as stochastic importation of rare strains maintains carriage of both sensitive and resistant strains of each serotype at low prevalence. The lowest-ranked eliminated serotypes (*e.g.* serotypes 25–30) exhibit less variability in resistance prevalence than the highest-ranked eliminated serotypes (*e.g.* serotypes 9–15) because the lower-ranked serotypes are more quickly cleared away when they do occasionally recirculate, meaning that the calculated resistance prevalence is more highly dominated by the fixed value of $Y_{\min}$. See Fig. 5, main text, for details.

---

Serotype-specific parameters for the extended mixed-carriage model run in Fig. 5 of the main text are given in **Appendix S4**. Introducing serotype-specific immunity to this model was necessary because serotype-specific clearance alone was insufficient to support the high diversity of pneumococcal serotype carriage observed in human populations, with only 8 of the 30 serotypes maintained (**Fig. S11**).

*Repeatability of model runs* — In Figs. 5 and 6 of the main text, we show results from various runs of the extended individual-based mixed-carriage model. Each plot summarises results from a single run rather than from multiple runs. To show that simulation results presented in the main text are repeatable, we show results from multiple independent runs here. Running the model multiple times necessitated using smaller population sizes and a coarser time step so that the runs would finish in a reasonable amount of time. This means that the trends are noisier than with larger population sizes, but the results are equivalent overall (**Fig. S12**).

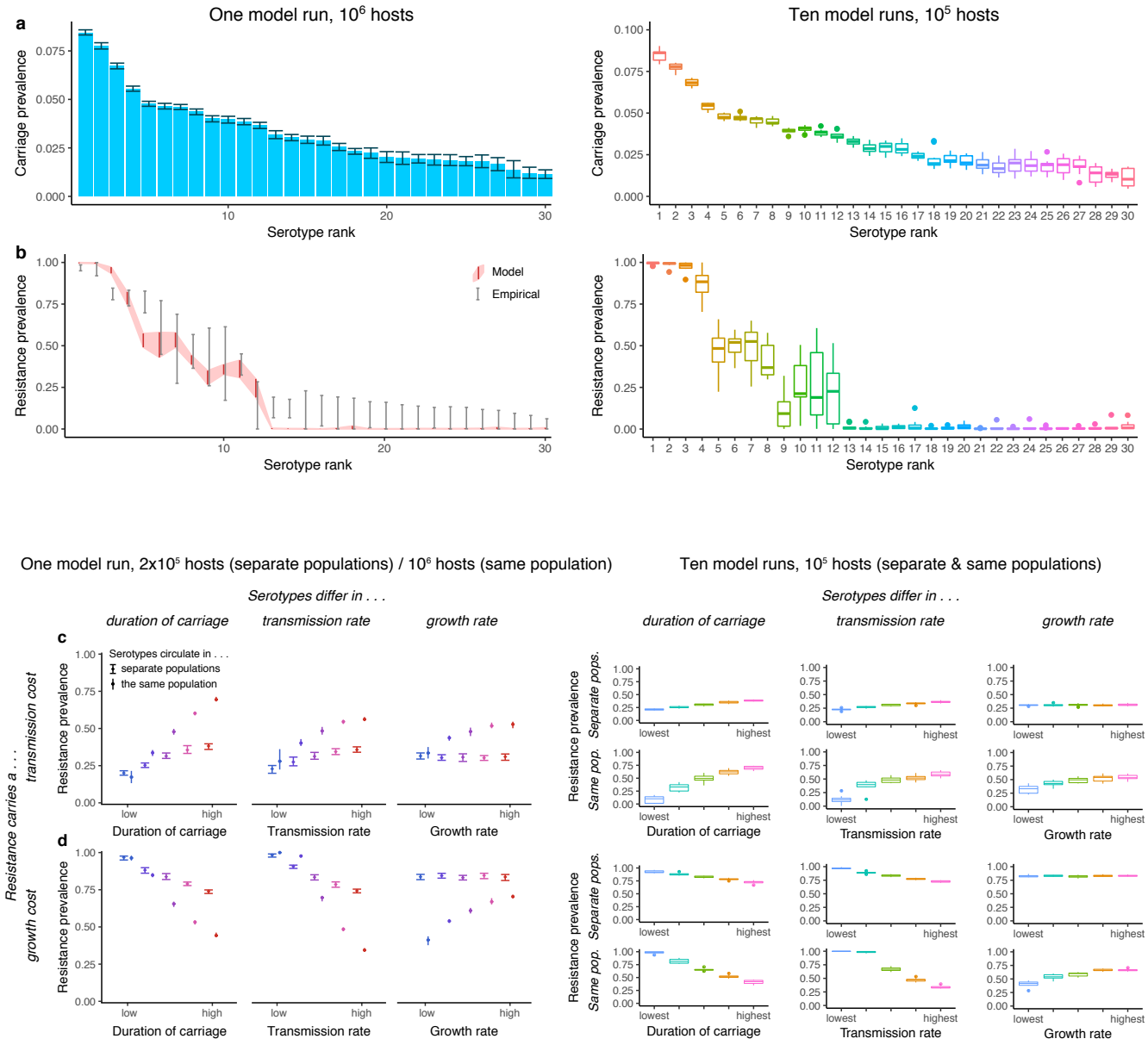



**Fig. S12 | Repeatability of the stochastic individual-based model implementation.** **(a, b)** The simulation which produced Fig. 5 of the main text (left) is repeated 10 times (*right*). Boxplots summarise the carriage prevalence and resistance prevalence of each serotype at simulation end over the 10 runs. These 10 runs used a smaller population size, $N = 10^5$, and a coarser time step, $\Delta t = 1/32$ mo$^{-1}$. See Fig. 5 (main text) for details. **(c, d)** The simulations which produced Fig. 6b&c of the main text (*left*) are repeated 10 times each (*right*). Boxplots summarise the resistance prevalence of each serotype at simulation end over the 10 runs. These 10 runs used a smaller population size, $N = 10^5$, and a coarser time step, $\Delta t = 1/32$ mo$^{-1}$. See Fig. 6 (main text) for details.

## Supplementary Note 6. Long-term trends in resistance prevalence

*Overview — This supplementary note analyses European trends in penicillin resistance prevalence in S. pneumoniae since 2007 in greater detail and argues that there is no evidence for a significant "lag" between drug consumption and drug resistance in this data set, suggesting that penicillin resistance in S. pneumoniae may be at equilibrium. This is consistent with observed coexistence between resistant and sensitive strains being a stable equilibrium, rather than a transient phase on the way to competitive exclusion.*

In Fig. 1d of the main text, we argue that observed intermediate resistance prevalences reflect stable coexistence between sensitive and resistant strains, rather than a transient phase on the way to competitive exclusion, because average resistance prevalence in the four pathogen-drug combinations we are investigating has essentially not changed from 2007–2015. The average resistance prevalence in Europe for 2007–2015 was calculated as a weighted mean of the resistance prevalence for each country[14,15]—with resistance prevalence sampled 1000 times from a beta distribution with parameters $\alpha = r + 1, \beta = n + 1 - r$ (*i.e.* assuming a uniform prior for the underlying binomial probability)—each time weighted by the population of the country in the corresponding year, across only those countries reporting resistance data for all years in 2007–2015, which left *ca.* 20 countries in each pathogen-drug data set.

Another way of looking at this question is to ask whether high consumption in a given year tends to predict a large increase in resistance in the following year. Looking at each European country in the data set from 2007–2015, there is a clear trend that penicillin consumption in a given year strongly predicts the percentage of *S. pneumoniae* isolates testing as non-susceptible in that year (**Fig. S13a**). However, having high consumption in a given year is not significantly associated with an increase in penicillin non-susceptibility into the next year (**Fig. S13b**) or two years hence (**Fig. S13c**). This seeming lack of a temporal relationship does not appear to be explained by countries with high consumption decreasing their consumption in further years, as there is no significant relationship between current consumption and the change in consumption in the following year (**Fig. S13d**). Taken together, this suggests that the response of resistant strains to antibiotic use may be relatively fast.

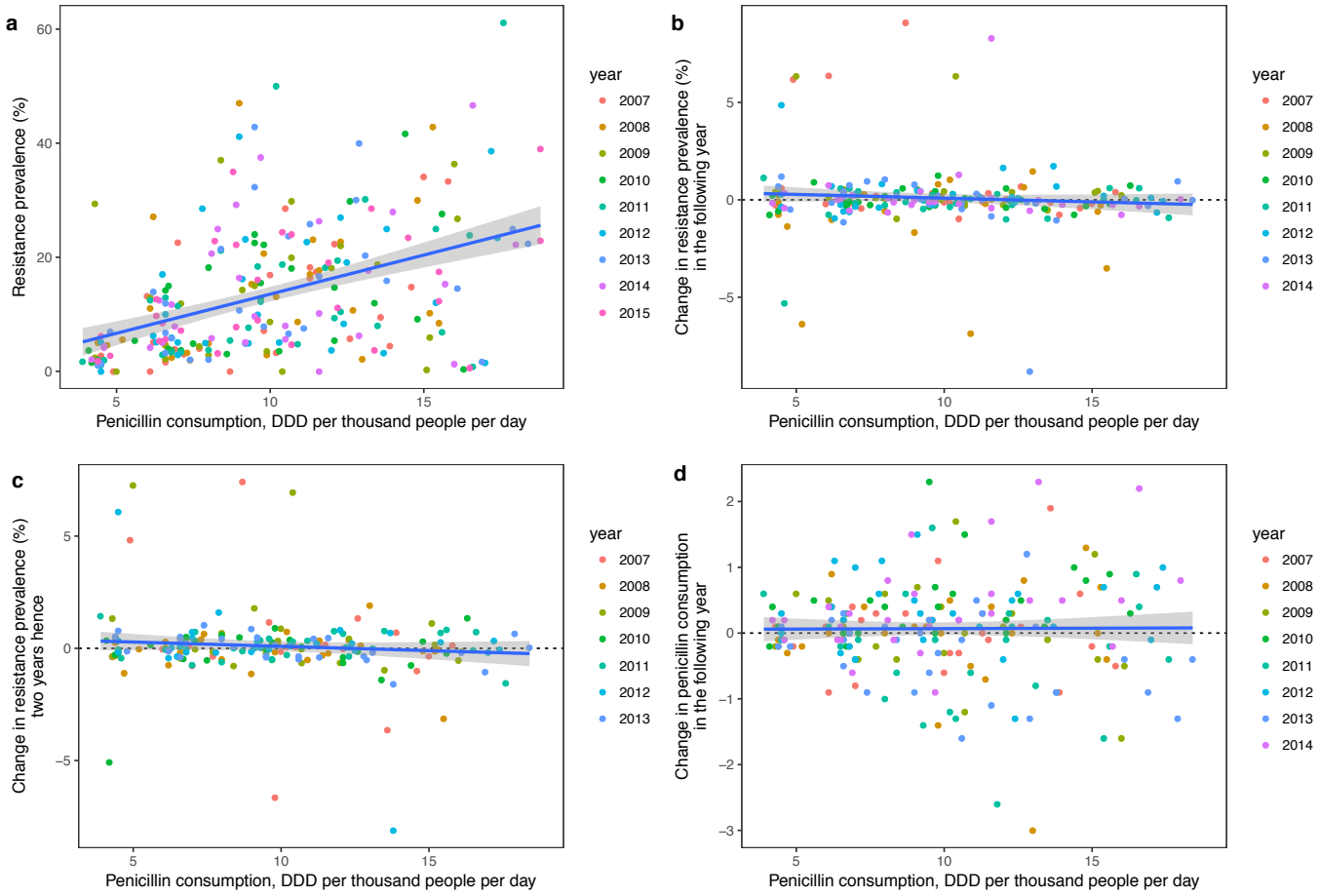



**Fig. S13 | Trends in resistance prevalence. (a)** Penicillin consumption predicts resistance prevalence in *S. pneumoniae* ($\beta = 0.443, F(1,250) = 61.07, P = 1.53\times10^{-13}$), but does not significantly predict either **(b)** the change from the current year to the next ($\beta = -0.0847, F(1,218) = 1.575, P = 0.211$), **(c)** the change between the current year and two years hence ($\beta = -0.140, F(1,190) = 3.801, P = 0.0527$), or **(d)** the change in consumption from the current year to the next ($\beta = 6.73\times10^{-3}, F(218,1) = 9.89\times10^{-3}, P = 0.921$). We report standardized coefficients ($\beta$), F-statistics and *P*-values for the slope term in a linear regression. Linear regressions are shown with 95% confidence intervals. DDD = defined daily doses.

## Supplementary Note 7. Data sources and interpretation of resistance

*Overview — This supplementary note summarizes the data sources used for analysis of drug resistance and drug consumption across European countries. In this section, we also defend our modelling assumption that invasive isolates drawn from carriers of both resistant and sensitive strains would not necessarily test positive for resistance.*

We summarize the data used and sources for the four data sets analysed (**Table S2**). The ECDC reports the number of invasive isolates that are susceptible (*i.e.* not resistant), intermediate (*i.e.* partially resistant), and resistant (*i.e.* highly resistant) out of all isolates tested. We decided to use the proportion of isolates that were intermediate or resistant (proportion non-susceptible) as the resistance prevalence of *S. pneumoniae*, and the proportion of isolates that were fully resistant (proportion resistant) as the resistance prevalence of *E. coli*, in keeping with both ECDC reporting conventions (typically, *S. pneumoniae* non-susceptibility and *E. coli* resistance are the "headline" figures reported for AMR in these pathogens) and previous studies[8,16] (which have focused on *S. pneumoniae* non-susceptibility).

*Interpretation of resistance* — Note that we assume that the overall frequency of resistant cells in the population is the appropriate proxy for resistance prevalence among invasive isolates in our models. That is, we do not count an individual who carries 1/2 resistant and 1/2 sensitive bacteria as being "clinically resistant", even though if one were to take a large sample of that individual's bacterial carriage, it would test positive for resistance. Rather, we assume that that individual, should they progress to an invasive disease state, has a 50% probability of yielding a resistant isolate.

| Data set | Consumption:<br>*ATC code, sector, year, and source* | Resistance:<br>*Pathogen, resistance metric, year, and source* |
|---|---|---|
| *E. coli* aminopenicillin resistance, 2015 (Fig. 4a, main text) | J01C (Beta-lactam antibacterials, penicillins), primary care, 2015[15] | *E. coli*, percentage resistant to aminopenicillins, 2015[17] |
| *E. coli* fluoroquinolone resistance, 2015 (Fig. 4b, main text) | J01MA (Fluoroquinolones), primary care, 2015[15] | *E. coli*, percentage resistant to fluoroquinolones, 2015[17] |
| *S. pneumoniae* macrolide resistance, 2015 (Fig. 4c, main text) | J01FA (Macrolides), primary care, 2015[15] | *S. pneumoniae*, percentage non-susceptible to macrolides, 2015[17] |
| *S. pneumoniae* penicillin resistance, 2007 (Fig. 4d, main text) | J01C (Beta-lactam antibacterials, penicillins), primary care, 2007[15,18†] | *S. pneumoniae*, percentage non-susceptible to penicillin, 2007[19] |

**Table S2.** Data sources for antibiotic consumption and antimicrobial resistance across five pathogen-drug combinations. †Portugal recorded no penicillin consumption for 2007 in the online ECDC database[15], but a 2011 ECDC report[18] provides the corrected figure of 11.3 defined daily doses per 1000 inhabitants per day for 2007.

This is because the data we are comparing our model output to is the fraction of invasive isolates which test positive for resistance, not the fraction of carriers who carry any resistant bacteria. Invasive disease is caused when a small number of (typically) genetically-identical cells leaves the normal, commensal host niche and invades the bloodstream or other normally-sterile sites. Isolates from blood or cerebrospinal fluid represent a sample of these invasive cells, and the protocol for testing resistance from these invasive cells involves isolating a single colony-forming unit from this sample. Accordingly, tested isolates are very likely to represent a single lineage of carried cells even when hosts carry multiple different strains.

Crucially, what we are modelling as carriage corresponds to commensal (non-invasive) carriage. Therefore, we make the assumption that if an individual carrying an equal number of resistant and sensitive bacteria progresses from commensal carriage to invasive disease, there will be a 50% chance (rather than a 100% chance) that an invasive isolate from the individual would test positive for resistance, and we assume that all carriers are equally likely to progress to invasive disease regardless of which strains they carry. Accordingly, the total fraction of invasive isolates testing positive for clinical resistance is equal to the total fraction of commensally-carried cells that are resistant, regardless of how these resistant cells are distributed among individual hosts.

## References


1. Leggett, H. C., Cornwallis, C. K. & West, S. A. Mechanisms of pathogenesis, infective dose and virulence in human parasites. *PLoS Pathog.* **8,** 10–12 (2012).
2. Colijn, C. *et al.* What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae*? *J. R. Soc. Interface* **7,** 905–919 (2010).
3. 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).
4. Kamng'ona, A. W. *et al.* High multiple carriage and emergence of *Streptococcus pneumoniae* vaccine serotype variants in Malawian children. *BMC Infect. Dis.* **15,** 234 (2015).
5. Turner, P. *et al.* Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *J. Clin. Microbiol.* **49,** 1784–1789 (2011).
6. Martinez-Medina, M. *et al.* Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* **15,** 872–882 (2009).
7. Mongkolrattanothai, K. *et al.* Simultaneous carriage of multiple genotypes of *Staphylococcus aureus* in children. *J. Med. Microbiol.* **60,** 317–322 (2011).
8. 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).
9. Li, B. *et al.* Duration of stool colonization in healthy medical students with extended-spectrum-β-lactamase-producing *Escherichia coli*. *Antimicrob. Agents Chemother.* **56,** 4558–4559 (2012).
10. Apisarnthanarak, A., Bailey, T. C. & Fraser, V. J. Duration of stool colonization in patients infected with extended-spectrum β-lactamase– producing *Escherichia coli* and *Klebsiella pneumoniae*. *Clin. Infect. Dis.* **46,** 1322–1323 (2008).
11. Ter Braak, C. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces. *Stat. Comput.* **16,** 239–249 (2006).
12. Plummer, M., Best, N., Cowles, K. & Vines, K. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* **6,** 7–11 (2006).
13. Cobey, S. & Lipsitch, M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science (80-. ).* **335,** 1376–1380 (2012).
14. European Centre for Disease Prevention and Control. Data from the ECDC Surveillance Atlas - Antimicrobial resistance. (2016). Available at: https://ecdc.europa.eu/en/antimicrobial-resistance/surveillance-and-disease-data/data-ecdc. (Accessed: 24th February 2018)
15. European Centre for Disease Prevention and Control. Antimicrobial consumption rates by country. (2018). Available at: http://ecdc.europa.eu/en/healthtopics/antimicrobial_resistance/esac-net-database/Pages/Antimicrobial-consumption-rates-by-country.aspx.
16. Goossens, H., Ferech, M., Vander Stichele, R. & Elseviers, M. Outpatient antibiotic use in Europe and association with resistance: A cross-national database study. *Lancet* **365,** 579–587 (2005).
17. European Centre for Disease Prevention and Control. *Antimicrobial resistance surveillance in Europe 2015. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. (ECDC, 2017).
18. European Centre for Disease Prevention and Control. *Surveillance of antimicrobial consumption in Europe*. (ECDC, 2014).
19. European Centre for Disease Prevention and Control. *EARSS Annual Report 2007*. (EARSS, 2007).

Appendix S1

**MCMC diagnostics from model fitting ($0 \leq k \leq 1$)**

Effective sample sizes (ESS) and upper bound of the 95% confidence interval for Gelman and Rubin's R ($R_{97.5}$) for MCMC. All effective sample sizes are > 1900 and all upper CIs for R are < 1.05. Calculated using the R package *coda.*

| | | $\beta$ | $c$ | $k$ | $u$ | $b$ |
|---|---|---|---|---|---|---|
| | | *E. coli* / Aminopenicillins | | | | |
| Knockout | ESS | 5340.47 | 5166.66 | 5372.3 | 5523.95 | N/A |
| | $R_{97.5}$ | 1.01 | 1.02 | 1.01 | 1.01 | N/A |
| Mixed-carriage, equal growth | ESS | 5137.58 | 4881.96 | 4732.2 | 5478.58 | N/A |
| | $R_{97.5}$ | 1.01 | 1.02 | 1.02 | 1.02 | N/A |
| Mixed-carriage, diff. growth | ESS | 5335.89 | N/A | 5076.31 | 5739.88 | 5809.44 |
| | $R_{97.5}$ | 1.02 | N/A | 1.02 | 1.02 | 1.02 |
| | | *E. coli* / Fluoroquinolones | | | | |
| Knockout | ESS | 6544.89 | 6526.19 | 6575.91 | 6466.13 | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | 1.01 | N/A |
| Mixed-carriage, equal growth | ESS | 5656.38 | 5605.33 | 4867.81 | 5904.84 | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.02 | 1.01 | N/A |
| Mixed-carriage, diff. growth | ESS | 5026.66 | N/A | 4422.71 | 5620.85 | 5589.87 |
| | $R_{97.5}$ | 1.01 | N/A | 1.02 | 1.01 | 1.01 |
| | | *S. pneumoniae* / Macrolides | | | | |
| Knockout | ESS | 8403.84 | 8150.76 | 8260.59 | N/A | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | N/A | N/A |
| Mixed-carriage, equal growth | ESS | 4877.06 | 1911.66 | 4591.86 | N/A | N/A |
| | $R_{97.5}$ | 1.02 | 1.02 | 1.01 | N/A | N/A |
| Mixed-carriage, diff. growth | ESS | 6334.32 | N/A | 6348.99 | N/A | 4190.04 |
| | $R_{97.5}$ | 1.01 | N/A | 1.01 | N/A | 1.03 |
| | | *S. pneumoniae* / Penicillins | | | | |
| Knockout | ESS | 5921.5 | 5893.09 | 5724.86 | N/A | N/A |
| | $R_{97.5}$ | 1.01 | 1.03 | 1.01 | N/A | N/A |
| Mixed-carriage, equal growth | ESS | 4763.84 | 4677.62 | 4719.38 | N/A | N/A |
| | $R_{97.5}$ | 1.02 | 1.01 | 1.01 | N/A | N/A |
| Mixed-carriage, diff. growth | ESS | 5128.53 | N/A | 5143.72 | N/A | 3728.18 |
| | $R_{97.5}$ | 1.02 | N/A | 1.01 | N/A | 1.02 |

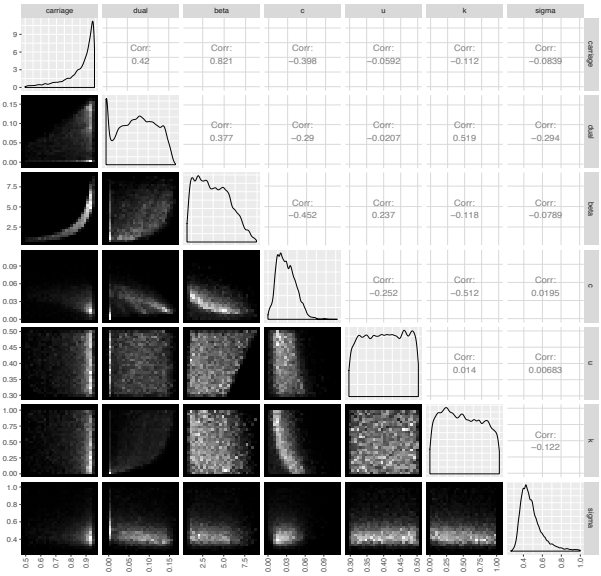
**MCMC diagnostics from model fitting ($0 \leq k \leq 5$)**

| | | $\beta$ | $c$ | $k$ | $u$ | $b$ |
|---|---|---|---|---|---|---|
| *E. coli* / Aminopenicillins | | | | | | |
| Knockout | ESS | 2561.41 | 2273.37 | 2515.04 | 2813.1 | N/A |
| | $R_{97.5}$ | 1.03 | 1.03 | 1.03 | 1.03 | N/A |
| Mixed-carriage, equal growth | ESS | 6085.44 | 5650.41 | 5796.76 | 5852.26 | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | 1.01 | N/A |
| Mixed-carriage, diff. growth | ESS | 5646.04 | N/A | 5702.31 | 6018.79 | 6011.58 |
| | $R_{97.5}$ | 1.01 | N/A | 1.01 | 1.01 | 1.01 |
| *E. coli* / Fluoroquinolones | | | | | | |
| Knockout | ESS | 1978.58 | 2011.53 | 2099.66 | 2597.95 | N/A |
| | $R_{97.5}$ | 1.03 | 1.04 | 1.03 | 1.04 | N/A |
| Mixed-carriage, equal growth | ESS | 5353.49 | 5506.28 | 5007.36 | 5614.77 | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.02 | 1.01 | N/A |
| Mixed-carriage, diff. growth | ESS | 5965.55 | N/A | 5744.64 | 6516.3 | 6407.98 |
| | $R_{97.5}$ | 1.02 | N/A | 1.01 | 1.01 | 1.01 |
| *S. pneumoniae* / Macrolides | | | | | | |
| Knockout | ESS | 7988.45 | 8283.92 | 8177.58 | N/A | N/A |
| | $R_{97.5}$ | 1 | 1.01 | 1 | N/A | N/A |
| Mixed-carriage, equal growth | ESS | 6790.96 | 6129.58 | 6825.4 | N/A | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | N/A | N/A |
| Mixed-carriage, diff. growth | ESS | 7267.47 | N/A | 7256.43 | N/A | 6644.38 |
| | $R_{97.5}$ | 1.01 | N/A | 1.01 | N/A | 1.01 |
| *S. pneumoniae* / Penicillins | | | | | | |
| Knockout | ESS | 3298.08 | 3155.58 | 3034.79 | N/A | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | N/A | N/A |
| Mixed-carriage, equal growth | ESS | 6618.19 | 6066.34 | 6529.98 | N/A | N/A |
| | $R_{97.5}$ | 1.01 | 1.01 | 1.01 | N/A | N/A |
| Mixed-carriage, diff. growth | ESS | 7462.49 | N/A | 7251.45 | N/A | 6144.1 |
| | $R_{97.5}$ | 1.01 | N/A | 1.01 | N/A | 1.02 |

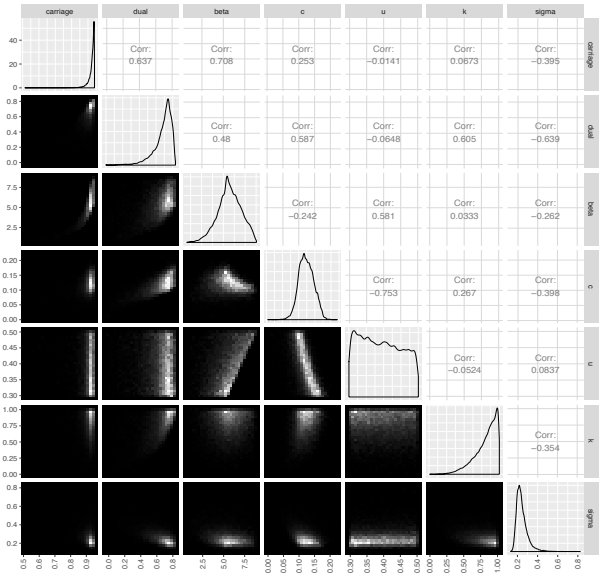
Appendix S2
**Joint posterior distributions**
*E. coli* / Aminopenicillins

These are pairwise joint posterior distributions and correlations between fitted parameters from model fitting. We focus on the case where $0 \leq k \leq 1$.

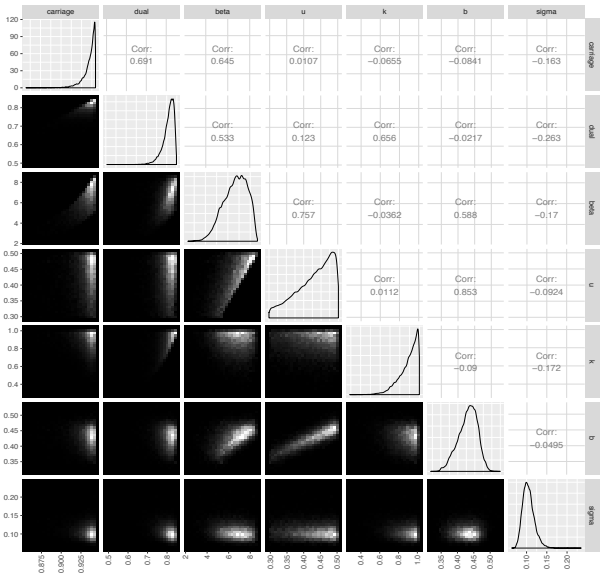
**Knockout** model




**Mixed-carriage** model (equal growth)

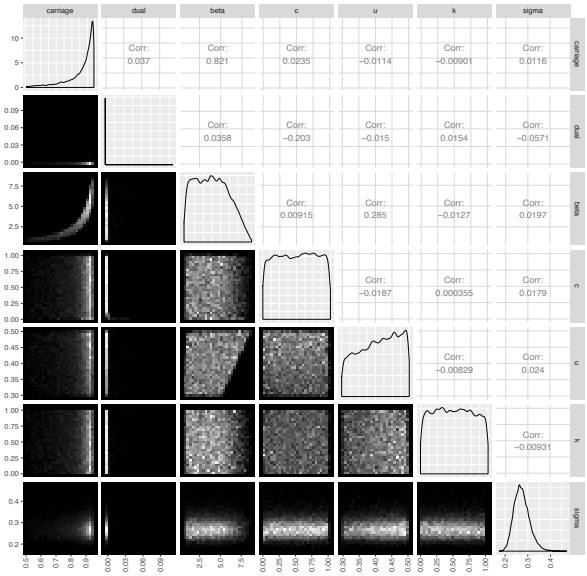



**Mixed-carriage** model (differential growth)

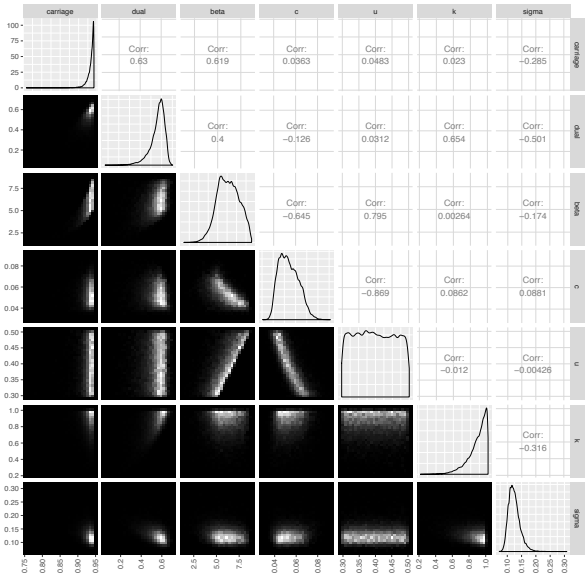


**Joint posterior distributions**
*E. coli* / Fluoroquinolones

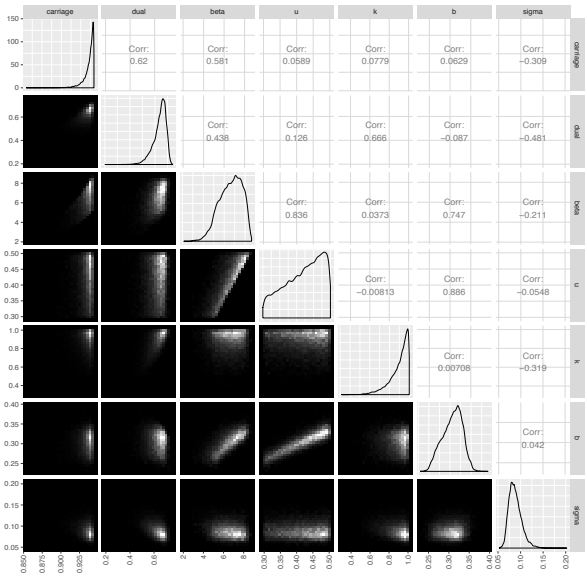
**Knockout**
model




**Mixed-carriage**
model
(equal growth)

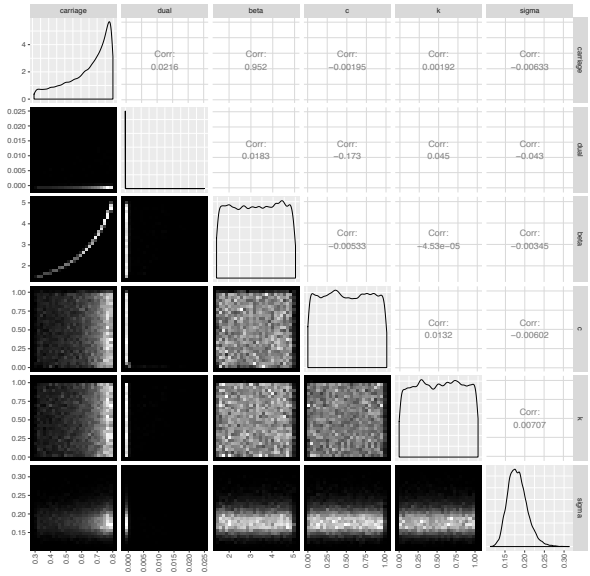



**Mixed-carriage**
model
(differential growth)

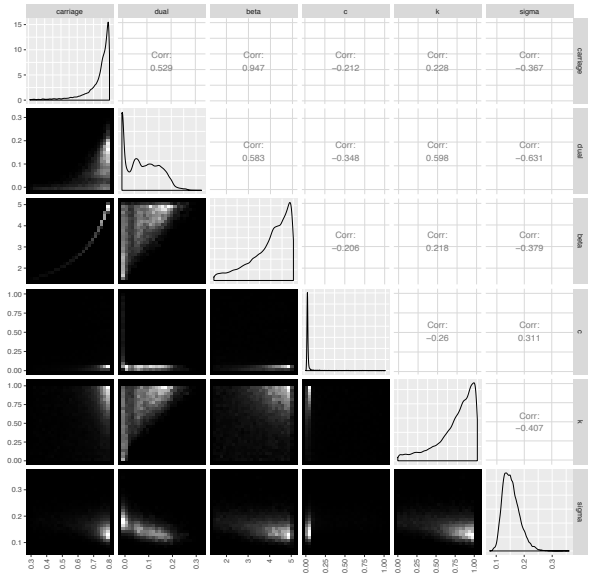


**Joint posterior distributions**
*S. pneumoniae* / Macrolides

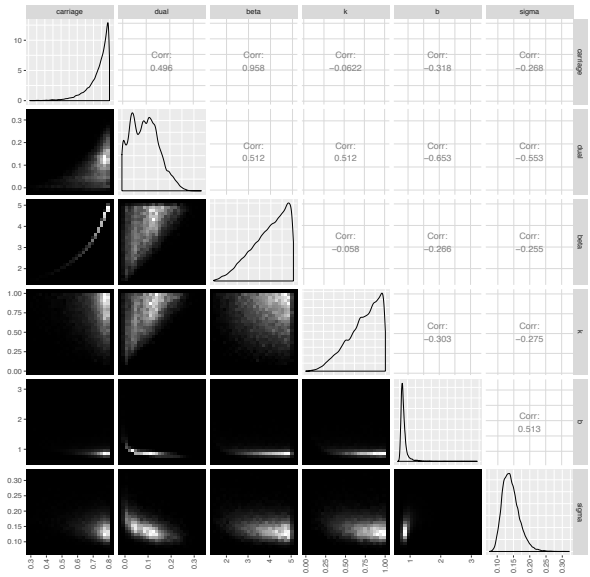
**Knockout**
model




**Mixed-carriage**
model
(equal growth)

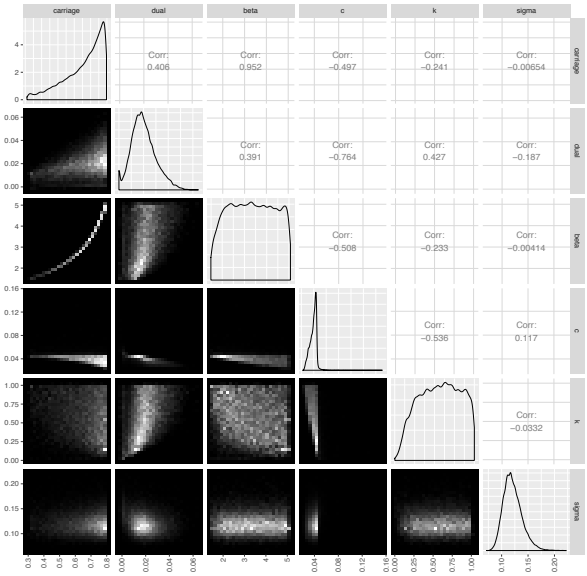



**Mixed-carriage**
model
(differential growth)

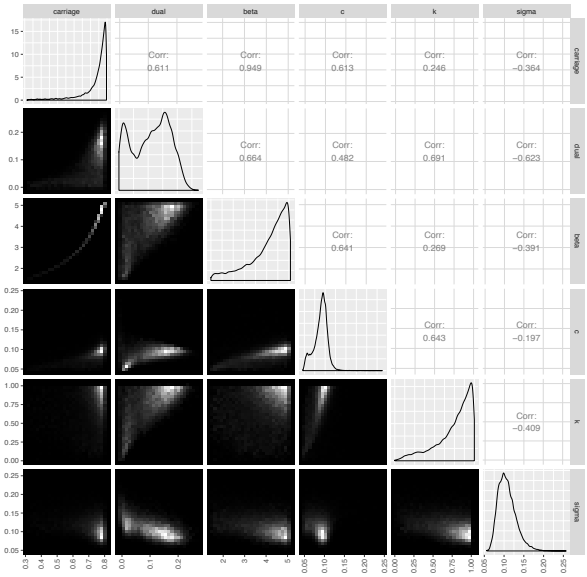


**Joint posterior distributions**
*S. pneumoniae* / Penicillin

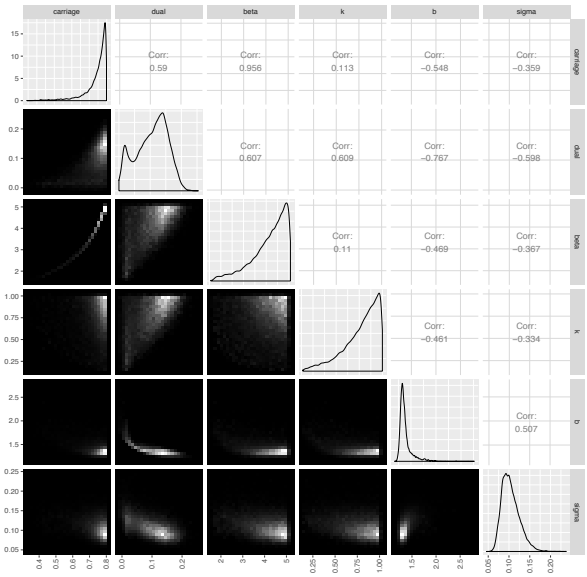
**Knockout**
model




**Mixed-carriage**
model
(equal growth)




**Mixed-carriage**
model
(differential growth)



Appendix S3

**Assessment of model fits**

We use AIC in the main text to formally assess model fit. Deviance, defined as $-2L$, where $L$ is the likelihood, is an alternative way of assessing model fit. Deviance provides a distribution rather than a single value, which some readers may prefer. We provide the AIC and the 95% HDI for deviance below, for both the case of $0 \le k \le 1$ (main text) and $0 \le k \le 5$ (Supplementary Note 2).

AIC

| | *E.coli* / Aminopenicillins | *E. coli* / Fluoroquinolones | *S. pneumoniae* / Macrolides | *S. pneumoniae* / Penicillins |
|---|---|---|---|---|
| MODEL | $0 \le k \le 1$ | | | |
| Knockout | 426.2 | 411.9 | 253.6 | 217.3 |
| Mixed-carriage *equal growth* | 414 | 387.5 | 247.8 | 215 |
| Mixed-carriage *differential growth* | 382.5 | 375.7 | 247 | 214.9 |
| | $0 \le k \le 5$ | | | |
| Knockout | 418.9 | 406.9 | 245.3 | 208.2 |
| Mixed-carriage *equal growth* | 399.5 | 370.7 | 234 | 206.2 |
| Mixed-carriage *differential growth* | 367.2 | 356.1 | 234 | 206.1 |

95% HDI for deviance

| | *E.coli* / Aminopenicillins | *E. coli* / Fluoroquinolones | *S. pneumoniae* / Macrolides | *S. pneumoniae* / Penicillins |
|---|---|---|---|---|
| MODEL | $0 \le k \le 1$ | | | |
| Knockout | 416.2–422.6 | 404.4–408.3 | 247.9–251.4 | 209.3–216.3 |
| Mixed-carriage *equal growth* | 404.2–413.9 | 377.7–386.6 | 239.9–249.4 | 207.0–216.4 |
| Mixed-carriage *differential growth* | 372.8–381.7 | 365.9–375.5 | 239.5–247.7 | 206.9–215.2 |
| | $0 \le k \le 5$ | | | |
| Knockout | 409.1–414.1 | 396.9–405.6 | 239.5–243.2 | 200.3–209.1 |
| Mixed-carriage *equal growth* | 390.0–397.7 | 360.9–368.9 | 226.1–233.7 | 198.2–205.2 |
| Mixed-carriage *differential growth* | 357.3–364.9 | 346.3–353.9 | 226.1–233.5 | 198.1–204.7 |

Appendix S4

**Model parameters for 30 pneumococcal serotypes**

Serotype-specific parameters used for the individual-based model runs parameterised with 30 serotypes. The serotype-specific clearance rates ($u$) are derived from an infant pneumococcal carriage study (see Lehtinen *et al.,* 2017, for data source).

| Serotype (fitness rank) | Sensitive strain | | Resistant strain | | $u$ |
|---|---|---|---|---|---|
| | $w$ | $\beta$ | $w$ | $\beta$ | |
| 1 | 30 | 3.2 | 24 | 2.88 | 0.218040621 |
| 2 | 29 | 3.2 | 23.2 | 2.88 | 0.228524919 |
| 3 | 28 | 3.2 | 22.4 | 2.88 | 0.24931694 |
| 4 | 27 | 3.2 | 21.6 | 2.88 | 0.281635802 |
| 5 | 26 | 3.2 | 20.8 | 2.88 | 0.31037415 |
| 6 | 25 | 3.2 | 20 | 2.88 | 0.313250944 |
| 7 | 24 | 3.2 | 19.2 | 2.88 | 0.313897489 |
| 8 | 23 | 3.2 | 18.4 | 2.88 | 0.3242715 |
| 9 | 22 | 3.2 | 17.6 | 2.88 | 0.34137673 |
| 10 | 21 | 3.2 | 16.8 | 2.88 | 0.34253003 |
| 11 | 20 | 3.2 | 16 | 2.88 | 0.346826302 |
| 12 | 19 | 3.2 | 15.2 | 2.88 | 0.35742264 |
| 13 | 18 | 3.2 | 14.4 | 2.88 | 0.392980189 |
| 14 | 17 | 3.2 | 13.6 | 2.88 | 0.403939796 |
| 15 | 16 | 3.2 | 12.8 | 2.88 | 0.412149955 |
| 16 | 15 | 3.2 | 12 | 2.88 | 0.414961346 |
| 17 | 14 | 3.2 | 11.2 | 2.88 | 0.44995069 |
| 18 | 13 | 3.2 | 10.4 | 2.88 | 0.479758149 |
| 19 | 12 | 3.2 | 9.6 | 2.88 | 0.491383953 |
| 20 | 11 | 3.2 | 8.8 | 2.88 | 0.527151935 |
| 21 | 10 | 3.2 | 8 | 2.88 | 0.535504695 |
| 22 | 9 | 3.2 | 7.2 | 2.88 | 0.540260509 |
| 23 | 8 | 3.2 | 6.4 | 2.88 | 0.55102657 |
| 24 | 7 | 3.2 | 5.6 | 2.88 | 0.553030303 |
| 25 | 6 | 3.2 | 4.8 | 2.88 | 0.561193112 |
| 26 | 5 | 3.2 | 4 | 2.88 | 0.565365551 |
| 27 | 4 | 3.2 | 3.2 | 2.88 | 0.603505291 |
| 28 | 3 | 3.2 | 2.4 | 2.88 | 0.686606471 |
| 29 | 2 | 3.2 | 1.6 | 2.88 | 0.749178982 |
| 30 | 1 | 3.2 | 0.8 | 2.88 | 0.777919864 |