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 $f_{\min} = Y_{\min} = 0$, $w_s = 1$, ι and Δt 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 ι 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_{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 δ_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 δ_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 Δt. 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 δ_i', which are arbitrary values infinitesimally close to zero but which may differ from δ_i. Since all δ_i and δ_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 Δt, 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

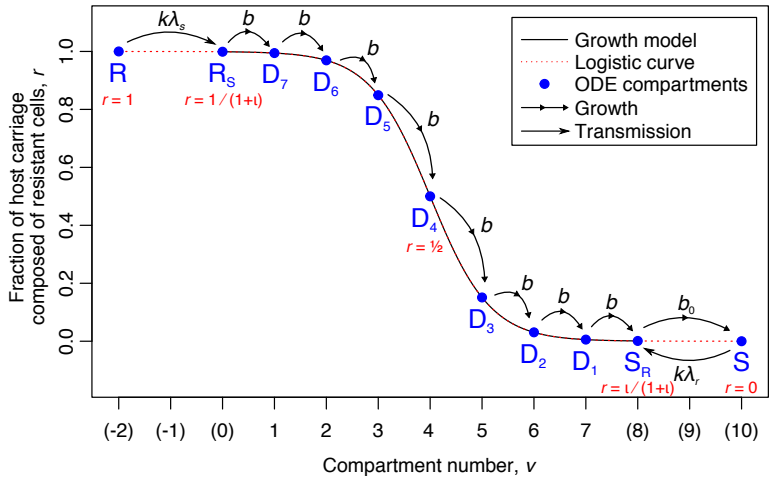


Fig. S1 | Schematic of within-host strain growth. Carriage compartments (filled blue circles) in the ODE approximation of the mixed-carriage model with differential within-host strain growth. Compartments between R_S and S_R are equally spaced along a logistic curve (dashed red line). The "growth model" 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 Δt. 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

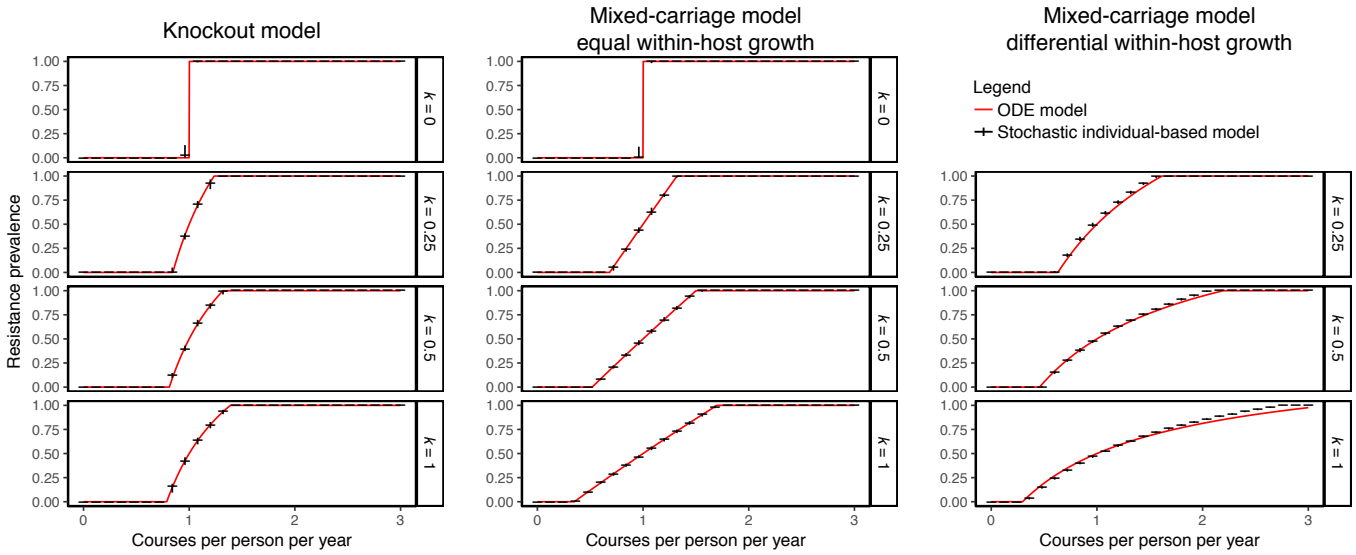


Fig. S2 | Correspondence between ODE and individual-based model implementations. We recapitulate the right-hand column of Fig. 3 of the main text to show that the individual-based implementations (black crosses) produce similar results to the ODE implementations (red lines) of each model. We assume $\iota = 0.001$, $f_{\min} = 3{\times}10^{-5}$, and N = 100,000. Across all models, we have $\beta = 5$ mo^{-1}, u = 1 mo^{-1}, and k and τ 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.

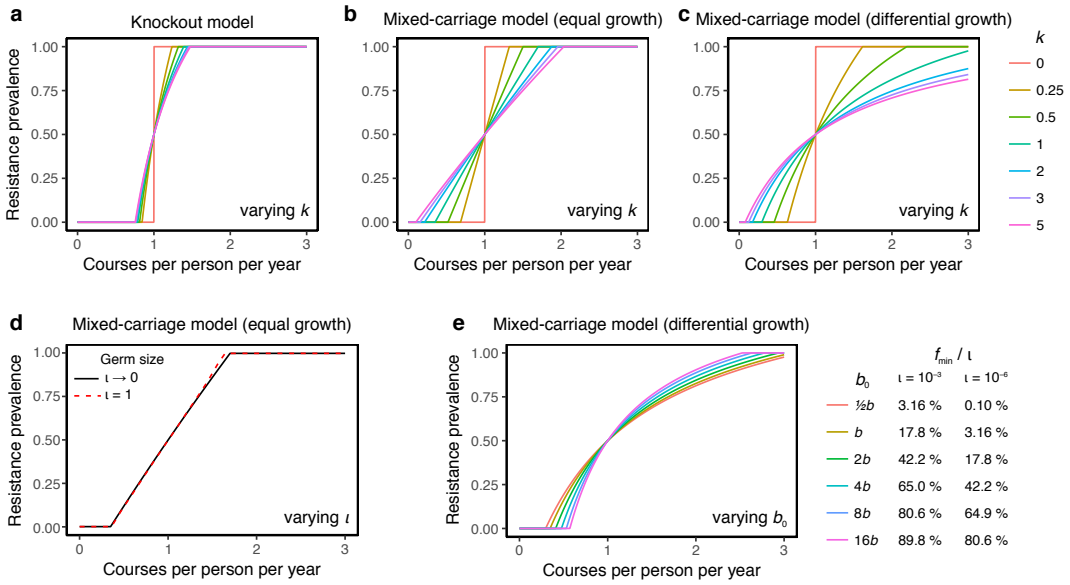


Fig. S3 | Factors inhibiting or promoting coexistence. (a–c) Increasing k above 1 promotes coexistence, but the effect is considerably smaller in the knockout model **(a)** than in the mixed-carriage models **(b, c)**. Parameters are otherwise the same as in Fig. 3c, f, i in the main text. **(d)** In 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 ι, 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

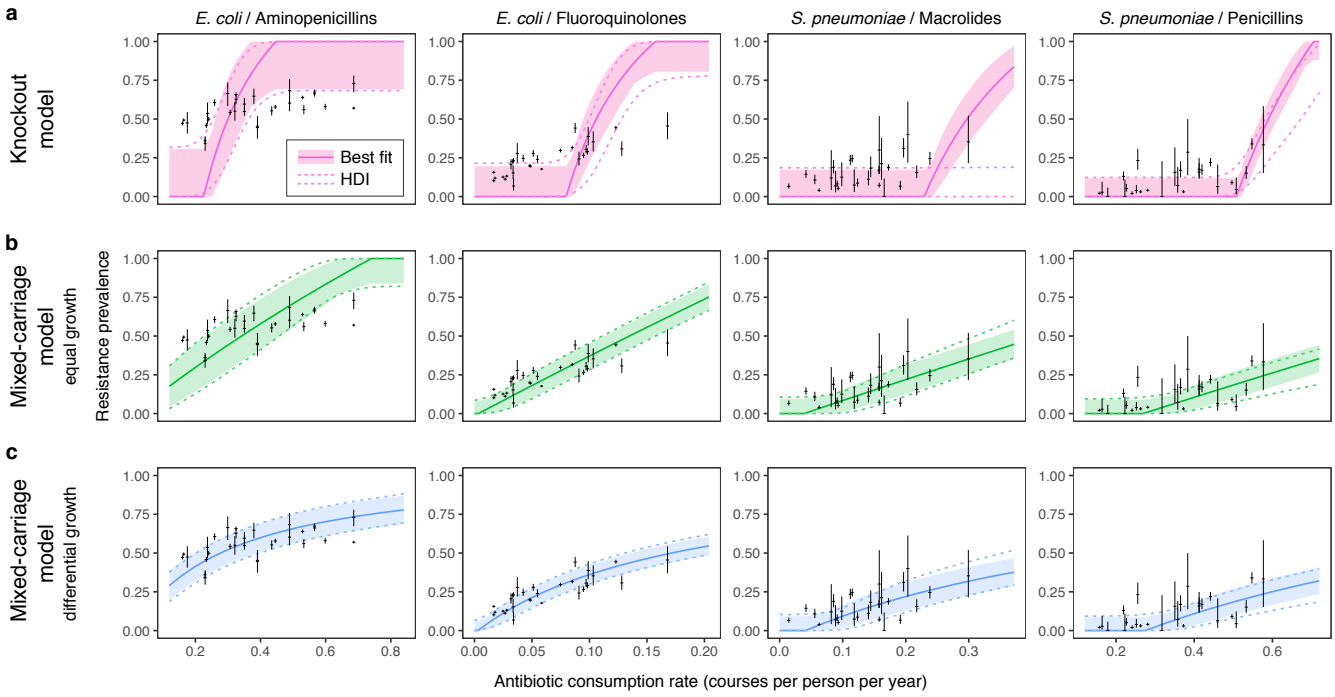


Fig. S4 | Model fits when we allow *k* to exceed 1 ($0 \leq k \leq 5$). Fits are not markedly different than when $0 \leq k \leq 1$ (see Fig. 4, main text). Solid lines and ribbons show the single best-fit run for each model (solid lines) and the 67% highest density interval incorporating between-country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

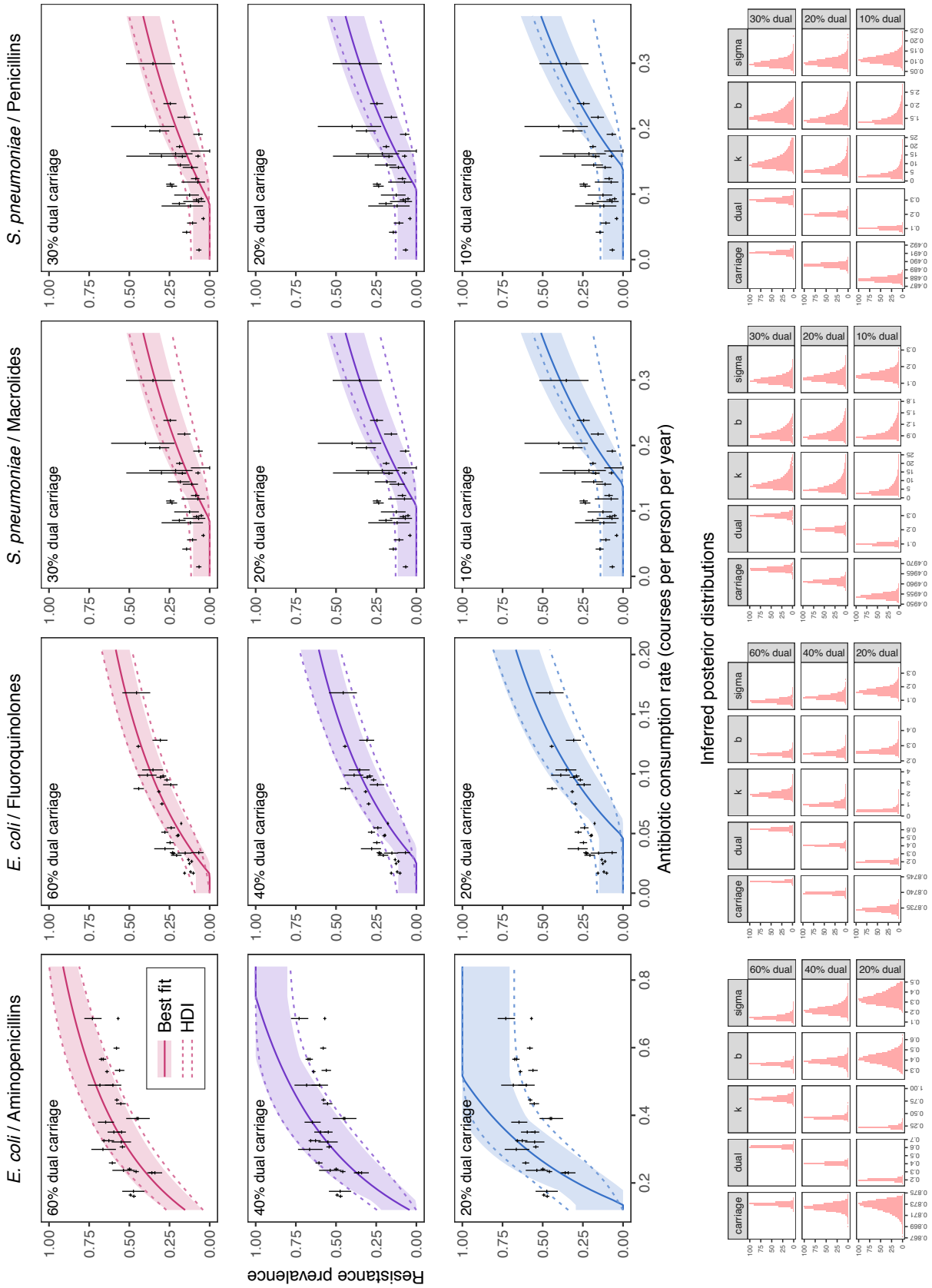


Fig. S5 | Model fits when dual carriage is constrained. Inferred posterior distributions are shown to illustrate constraints on dual carriage (panels marked "dual") and how other parameters change due to constraints on dual carriage: "k" and "sigma" show larger relative shifts, particularly for *E. coli*,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

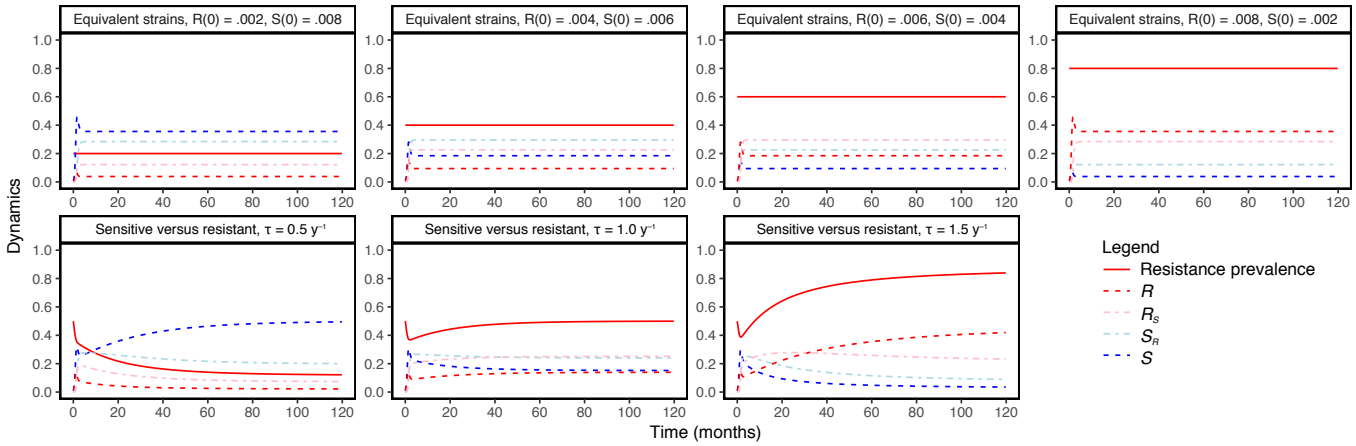


Fig. S6 | Dynamics of the ODE implementation of the mixed-carriage model. *Top row*: Model dynamics for the mixed-carriage ODE model where the S strain and R strain are equivalent. That is, we have $\tau = 0,\ c = 0$, and $b = 0$. Other parameters are $\beta = 5,\ u = 1$, and $k = 1$. Each panel of the 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 τ 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 β is indirectly constrained by a likelihood penalty on prevalence of carriage, so we set a uniform prior for β wide enough to overlap the full range of permissible carriage prevalence, Y, given the range of clearance rates u and treatment rates τ (*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			
β (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})
σ (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.

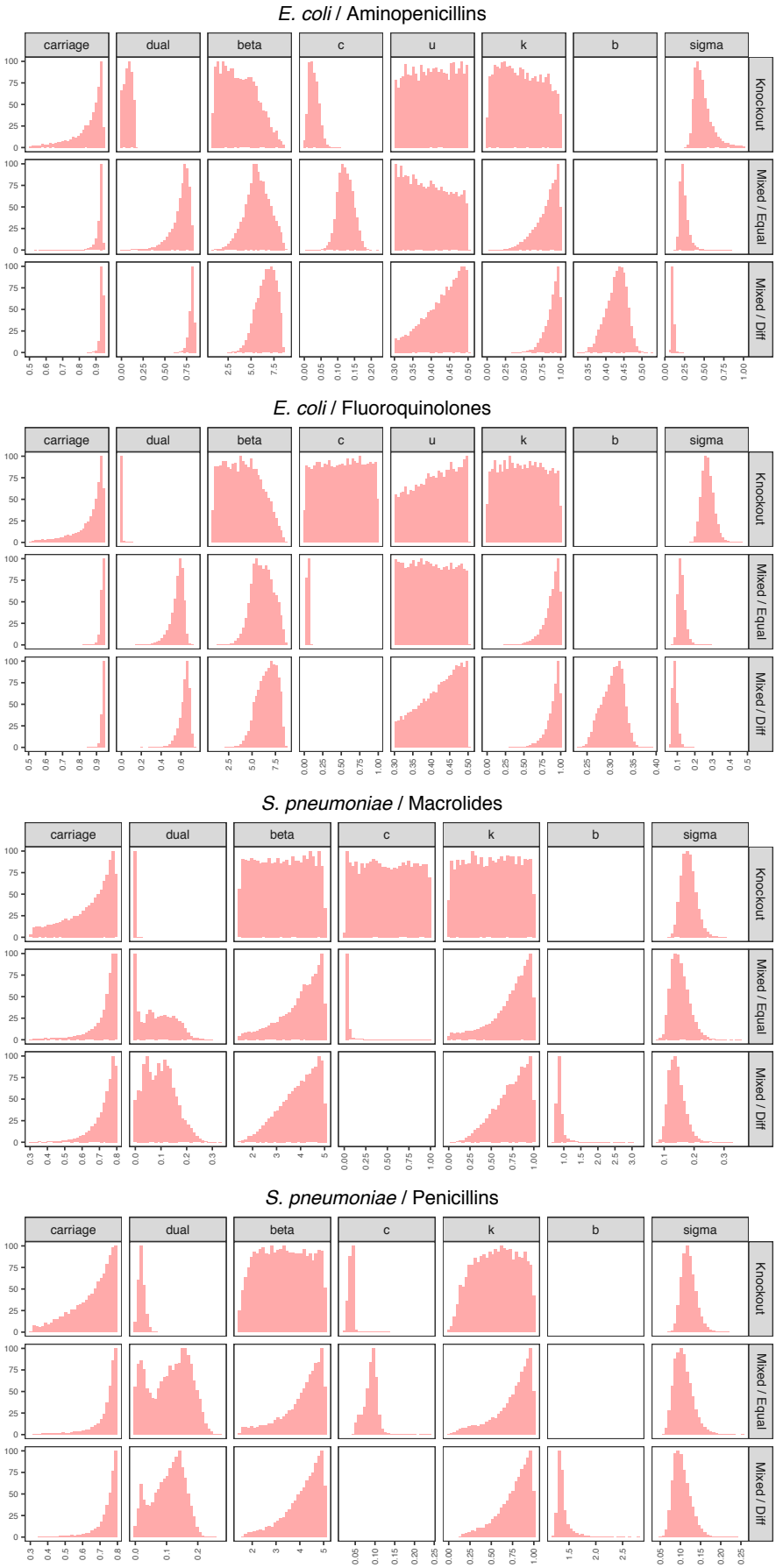


Fig. S7 | See previous page for caption.

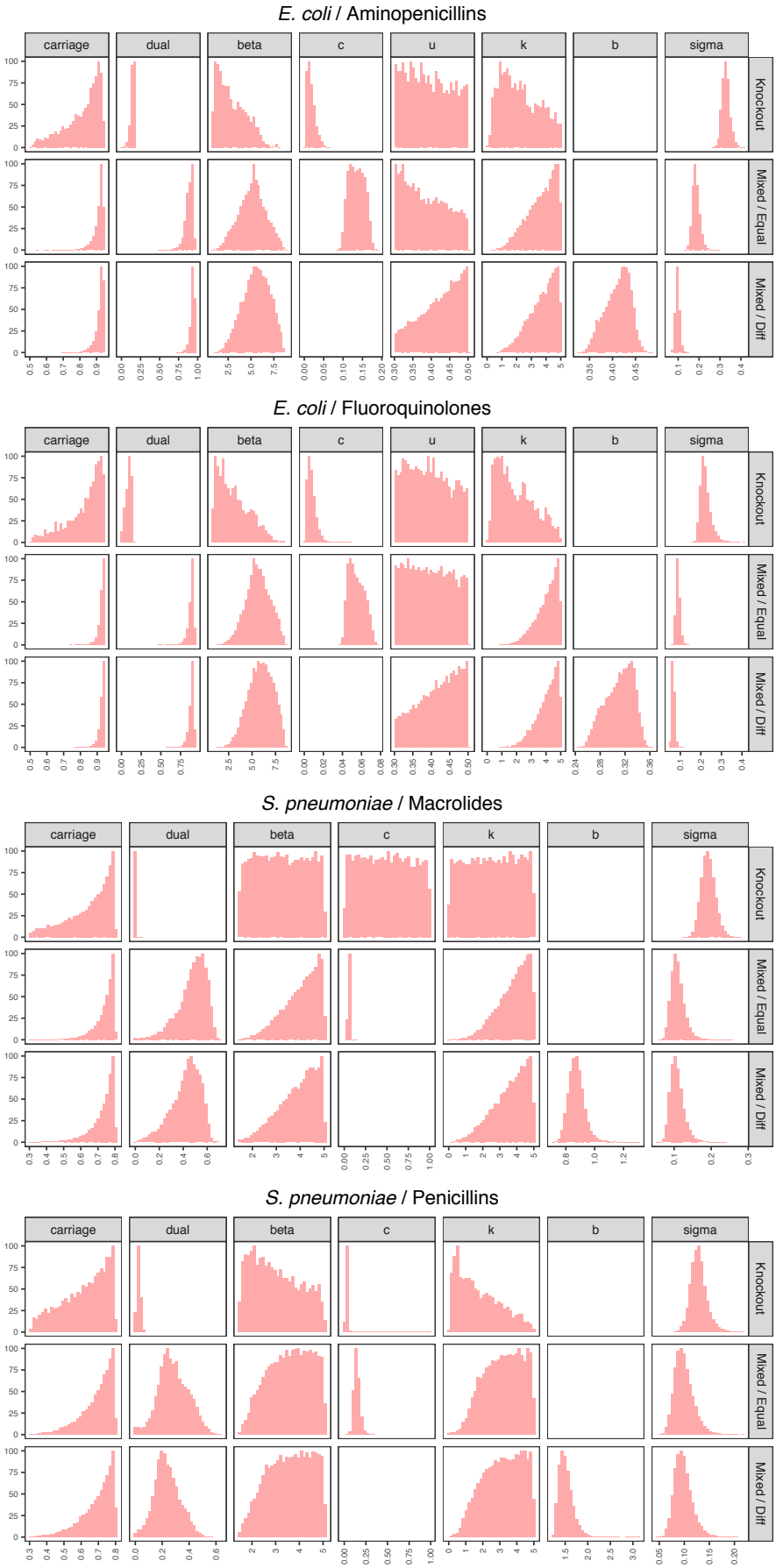


Fig. S8 | Posteriors for model fitting when $0 \leq k \leq 5$. See Fig. S7 for details.

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, β_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 ℓ occur at rate u_ℓ to each host, where u_ℓ is the clearance rate for serotype ℓ; and

treatment events occur at rate τ to each host, where τ 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 ℓ 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 α (*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 ℓ and $m_{i,\ell} = 0$ if host *i* is not immune to serotype ℓ. 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 ℓ is blocked if the host is immune to that serotype; clearance of serotype ℓ by host *i* makes host *i* immune to serotype ℓ 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 $m_{i,\ell} = 0$.

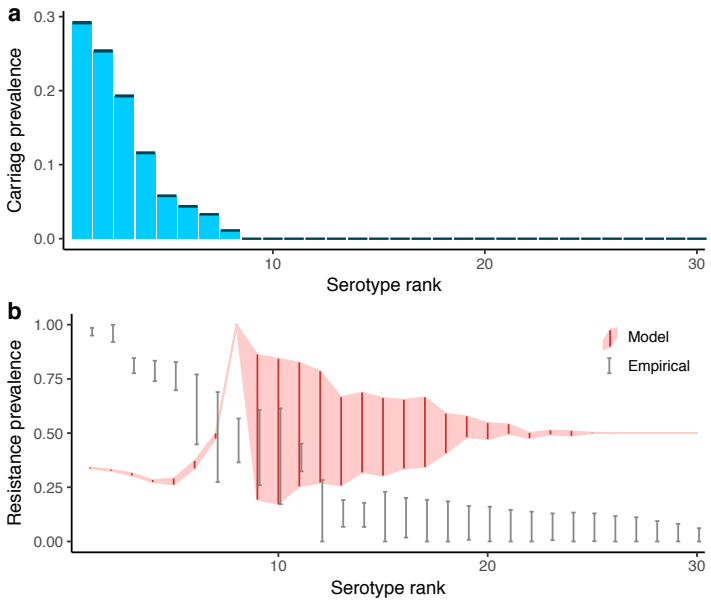


Fig. S11 | Immunity is needed to maintain significant serotype diversity. As serotypes vary greatly in duration of carriage, the mechanism of serotype-specific clearance alone is not able to reproduce observed patterns of pneumococcal carriage or resistance prevalence, as all but the 8 serotypes with 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**).

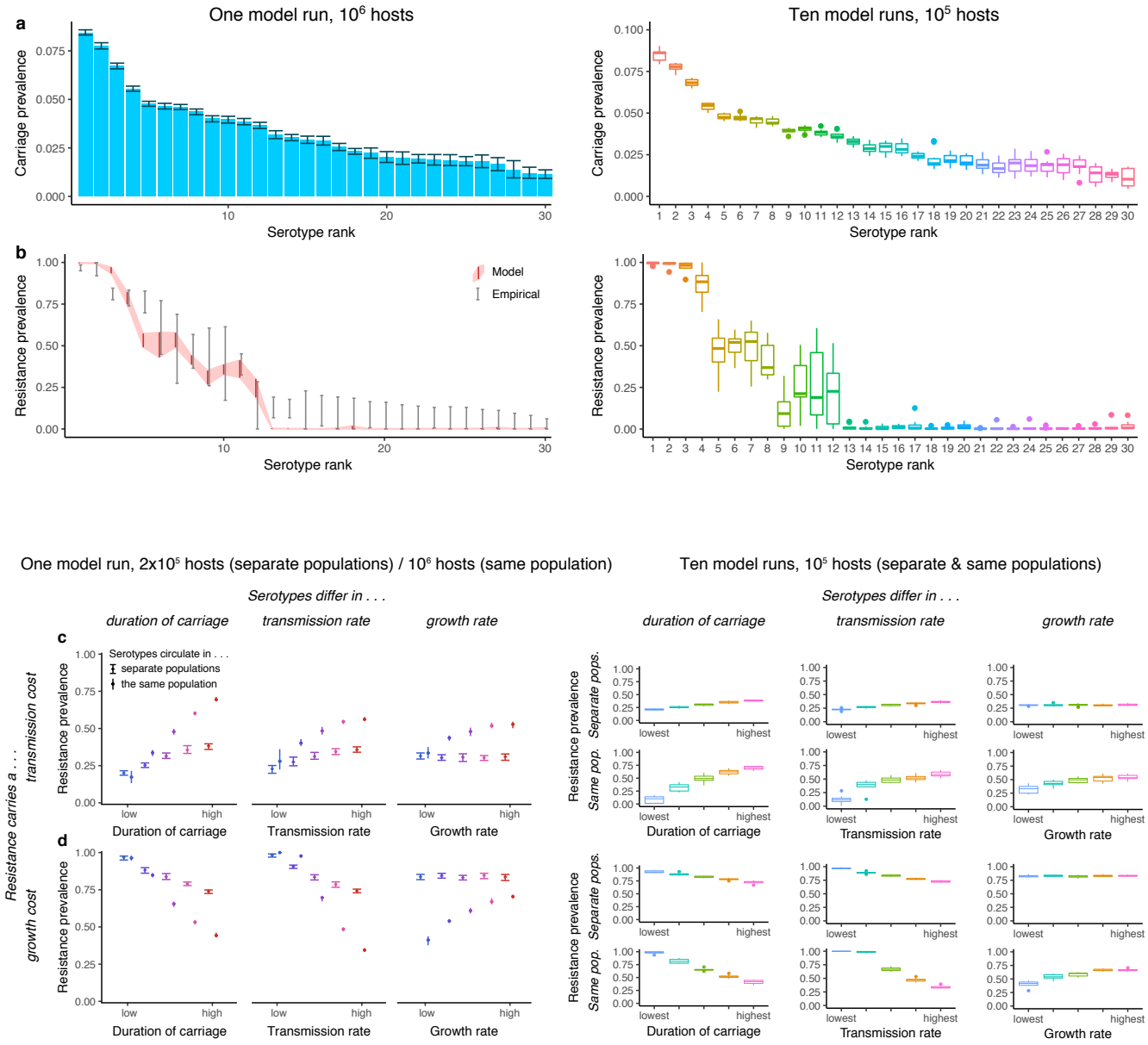


Fig. S12 | Repeatability of the stochastic individual-based model implementation. **(a, b)** The simulation which produced Fig. 5 of the main text (left) is repeated 10 times (*right*). Boxplots summarise the carriage prevalence and resistance prevalence of each serotype at simulation end over the 10 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.

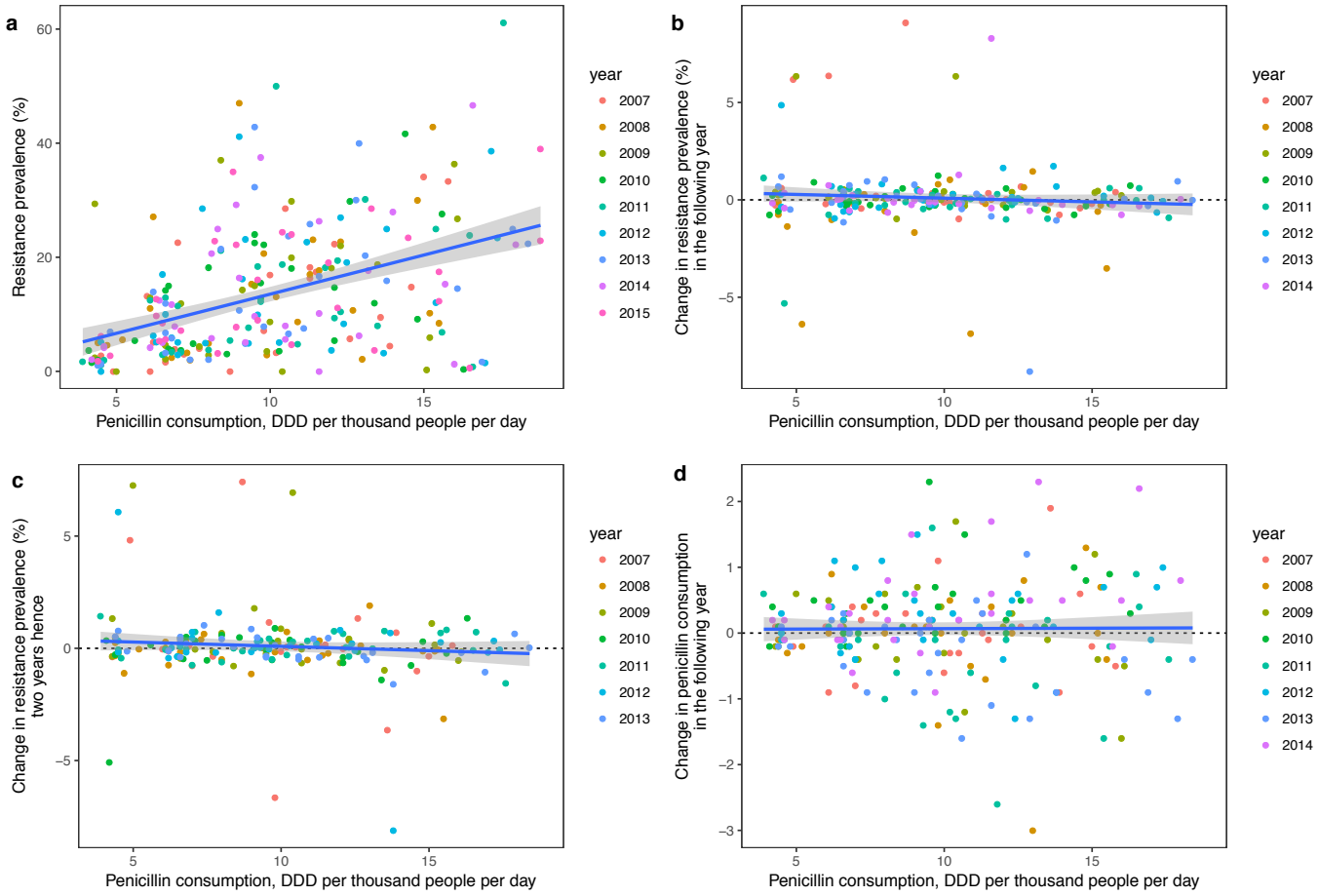


Fig. S13 | Trends in resistance prevalence. (a) Penicillin consumption predicts resistance prevalence in *S. pneumoniae* ($\beta = 0.443, F(1,250) = 61.07, P = 1.53\times10^{-13}$), but does not significantly predict either **(b)** the change from the current year to the next ($\beta = -0.0847, F(1,218) = 1.575, P = 0.211$), **(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 (β), 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: *ATC code, sector, year, and source*	Resistance: *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.*

		β	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

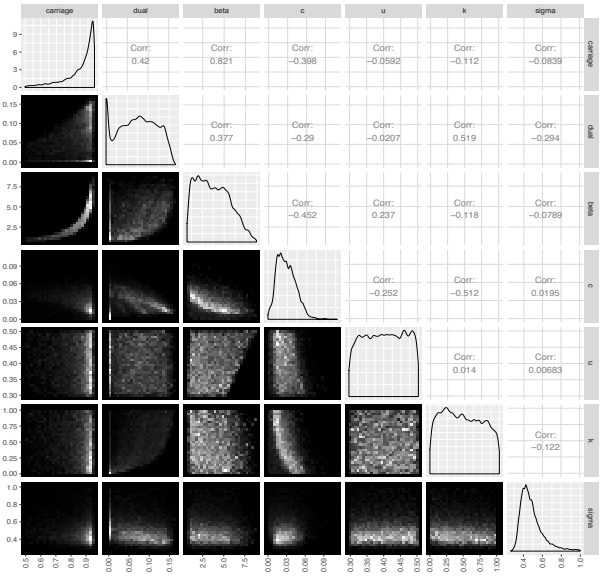
MCMC diagnostics from model fitting ($0 \leq k \leq 5$)

		β	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

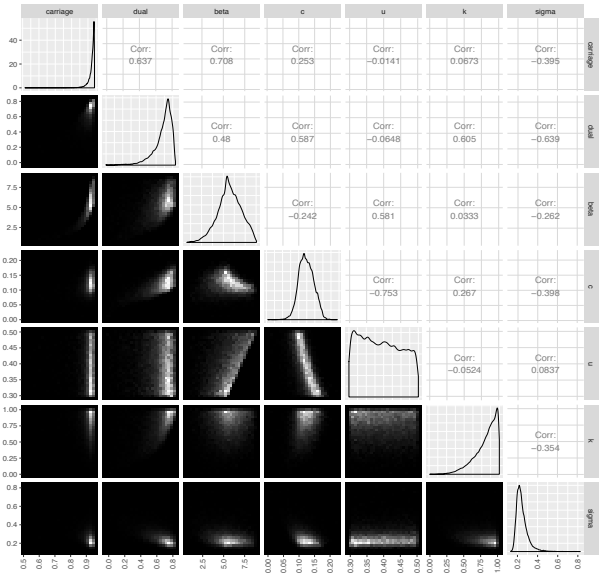
Appendix S2
Joint posterior distributions
E. coli / Aminopenicillins

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

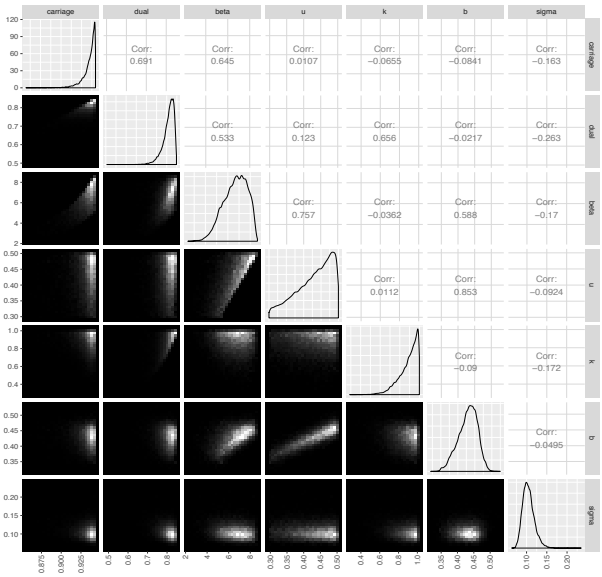
Knockout model



Mixed-carriage model (equal growth)

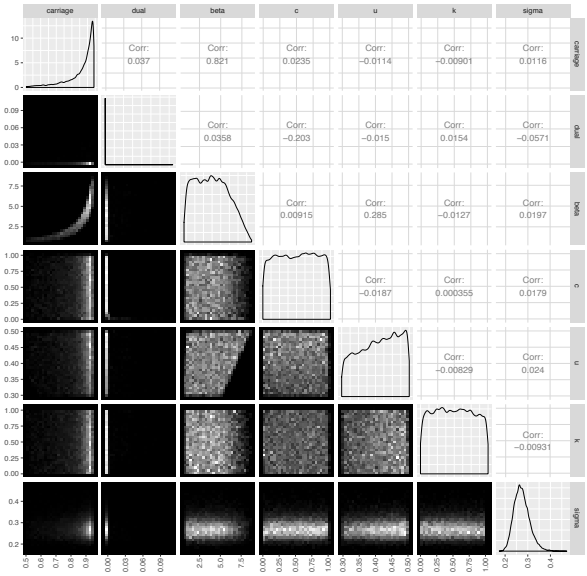


Mixed-carriage model (differential growth)

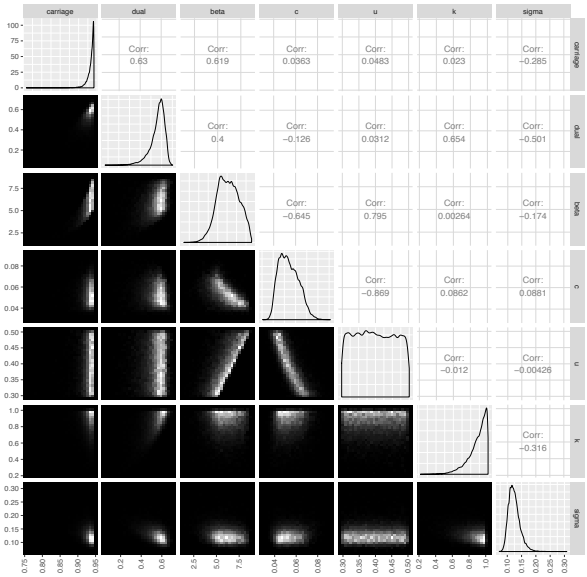


Joint posterior distributions
E. coli / Fluoroquinolones

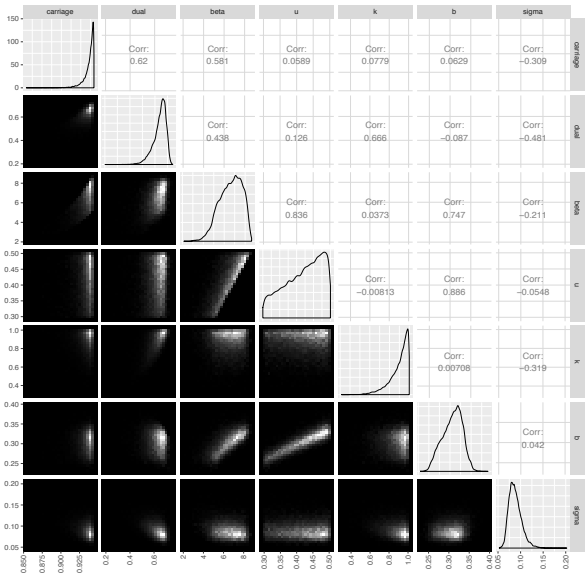
Knockout
model



Mixed-carriage
model
(equal growth)

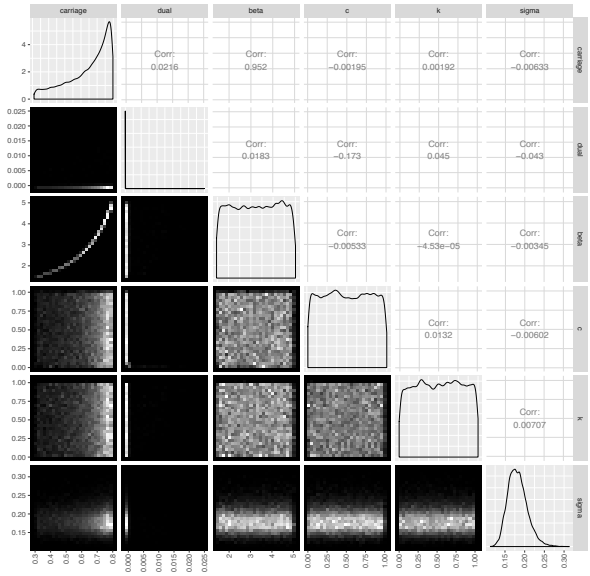


Mixed-carriage
model
(differential growth)

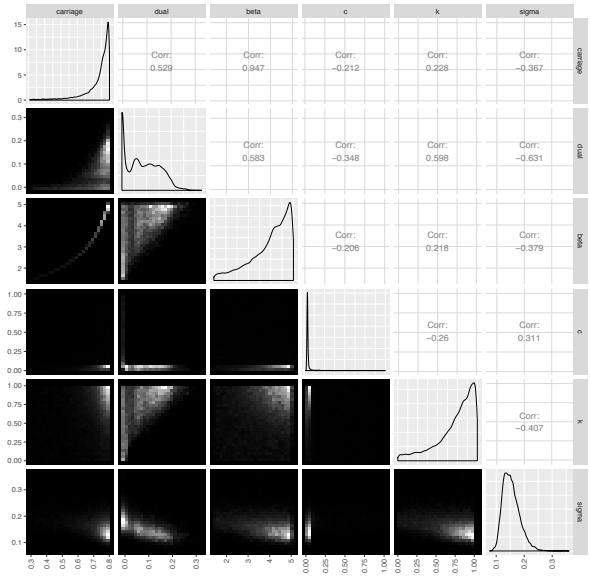


Joint posterior distributions
S. pneumoniae / Macrolides

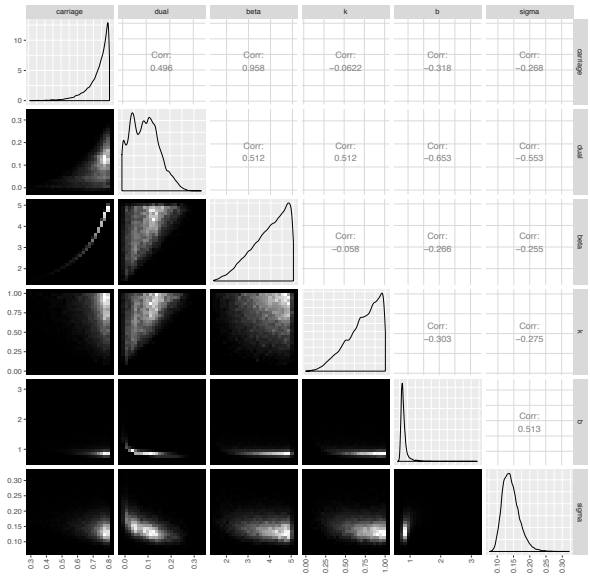
Knockout
model



Mixed-carriage
model
(equal growth)

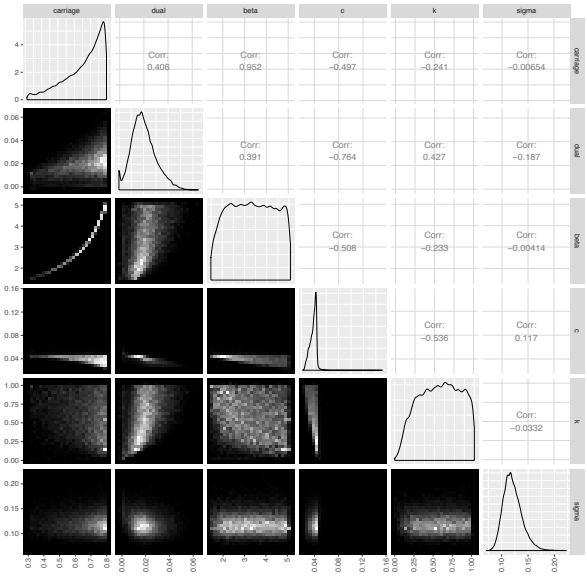


Mixed-carriage
model
(differential growth)

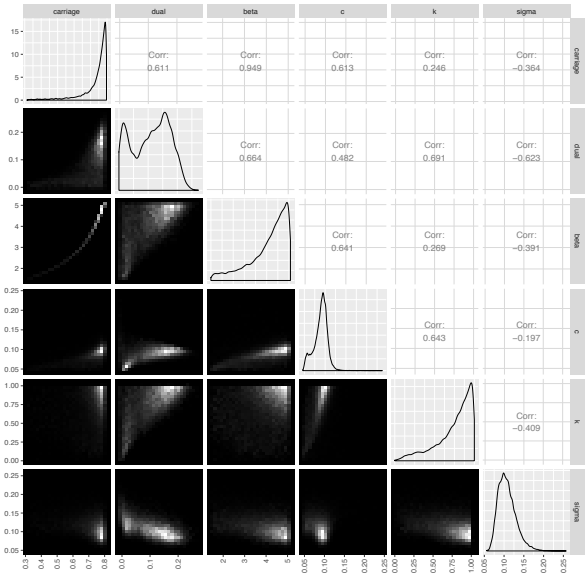


Joint posterior distributions
S. pneumoniae / Penicillin

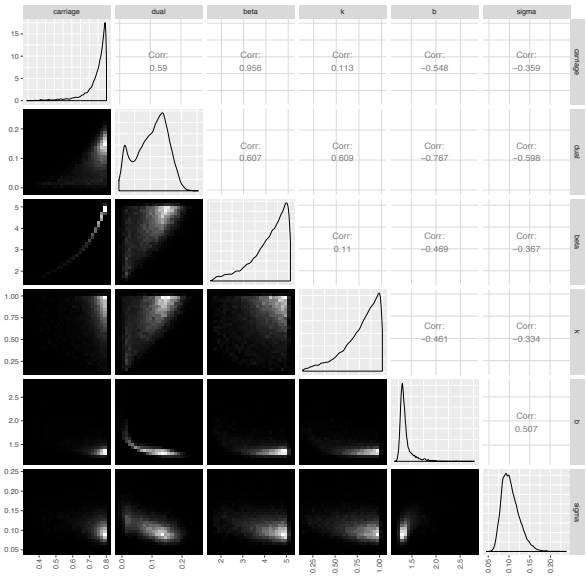
Knockout
model



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	β	w	β	
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