Quantifying the effects of antibiotic resistance and within-host competition on strain fitness in *Streptococcus pneumoniae*

Aswin Krishna [1], Gerry Tonkin-Hill [2], Thibaut Morel-Journel [3], Stephen Bentley [4], Paul Turner [5], François Blanquart [6], and Sonja Lehtinen [1,7,✉]

[1]Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zürich, Switzerland; [2]Peter Doherty Institute for Infection and Immunity, Melbourne, Australia; [3]Université Sorbonne Paris Nord, Université Paris Cité, INSERM, IAME, Bobigny 93000, France; [4]Wellcome Sanger Institute, Cambridge, UK; [5]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK; [6]Center for Interdisciplinary Research in Biology, CNRS, Collège de France, PSL Research University, Paris 75005, France; [7]Department of Computational Biology, University of Lausanne, Switzerland

Competition plays a key role in shaping the structure and diversity of bacterial populations. In many clinically important bacterial species, strains compete at multiple scales: at the between-host scale for new hosts to colonise, and at the within-host scale during co-colonisation. Characterising these multiple facets of competition plays an important role in understanding bacterial ecology. This is particularly relevant for antibiotic resistance, where competition between antibiotic-susceptible and resistant strains determines resistance dynamics. In this work, we perform survival analyses on a large longitudinal dataset of *Streptococcus pneumoniae* carriage to quantify how within-host competition affects the rates of clearance and establishment of pneumococcal strains. We find that the presence of a within-host competitor is associated with a 33% increase in clearance and a 54% reduction in establishment. Priority effects and serotype differences partially predict the outcomes of this within-host competition. Further, we quantify the effects of antibiotic resistance on between- and within-host components of fitness. Antibiotic consumption is associated with increased clearance rate for both susceptible and resistant strains, albeit to a higher extent in susceptible strains. In the absence of antibiotics, we find some evidence that resistance is associated with increased susceptibility to within-host competition, suggesting a fitness cost of resistance. Overall, our work provides quantitative insights into pneumococcal competition across scales and the role of this competition in shaping pneumococcal epidemiology.

Within-host competition | *Streptococcus pneumoniae* | Antibiotic resistance | Fitness cost of resistance

Correspondence: ***sonja.lehtinen@unil.ch***

Introduction

Several clinically important bacteria such as *Escherichia coli* and *Streptococcus pneumoniae* exhibit remarkable strain-level diversity. Strains of these species differ in key traits like virulence, mobility, metabolism or antibiotic-resistance (1); nevertheless, different strains can have large overlaps in their genomic content and niche occupancy (2, 3), resulting in significant intra-species competition. Understanding how these strains compete is important since competition is a major force shaping the observed distribution of strains in a species (4, 5).

Strains of a species compete with each other across multiple scales. While competition is conceived and modelled in many studies in terms of competition for acquisition of uncolonised hosts (6–10), strains can also compete for growth and survival within a host (11). In species like *E. coli* and *S. pneumoniae* where co-colonisation is common (e.g. median number of co-colonisers is around 1.5 in pneumococcal carriers (12) and 3 in *E. coli* carriers (13)), within-host competition potentially plays a major role in shaping strain dynamics. However, the specific effects of within-host competition on strain fitness at the epidemiological scale–e.g. its impact on duration of carriage, transmission to other hosts or the efficiency of colonisation–are not well understood. Previous modelling work has demonstrated that assumptions about these effects have significant impacts on population dynamics (14–21), highlighting the importance of this knowledge gap.

In this work, we study competition by quantifying its effects on the ability of strains to establish and persist in hosts (Figure 1). We focus on *S. pneumoniae*, a mostly asymptomatic coloniser of the nasopharynx that can become pathogenic when it evades host immune responses (22). *S. pneumoniae* is associated with a substantial disease burden in children and geriatric patients (23, 24) and features among the World Health Organization (WHO) list of priority pathogens for antibiotic resistance (25). Our focus is in particular on two aspects of pneumococcal competition: competition between serotypes (i.e. antigenic strain types defined based on the bacteria's polysaccharide capsule) and competition between antibiotic susceptible and resistant strains. Serotype competition is important because it determines how pneumococcal populations respond to vaccinations covering a subset of serotypes. The removal of

How does this ecological factor... affects ... affect these components of fitness?	Clearance rate	Establishment rate	Within-host competitive ability
Presence of competitors	*Increases*	*Decreases*	*Not applicable*
Cost of antibiotic resistance	*Confounded by association of resistance with low clearance rate*	*Not testable in this data*	*Decreases for some antibiotics*

Figure 1. In this study, we are interested in how the presence of within-host competitors and cost of antibiotic resistance shapes the different components of fitness of *S. pneumoniae* strains. The table here summarizes the evidence we find for associations between each ecological factor (in columns) and the fitness measures (in rows).

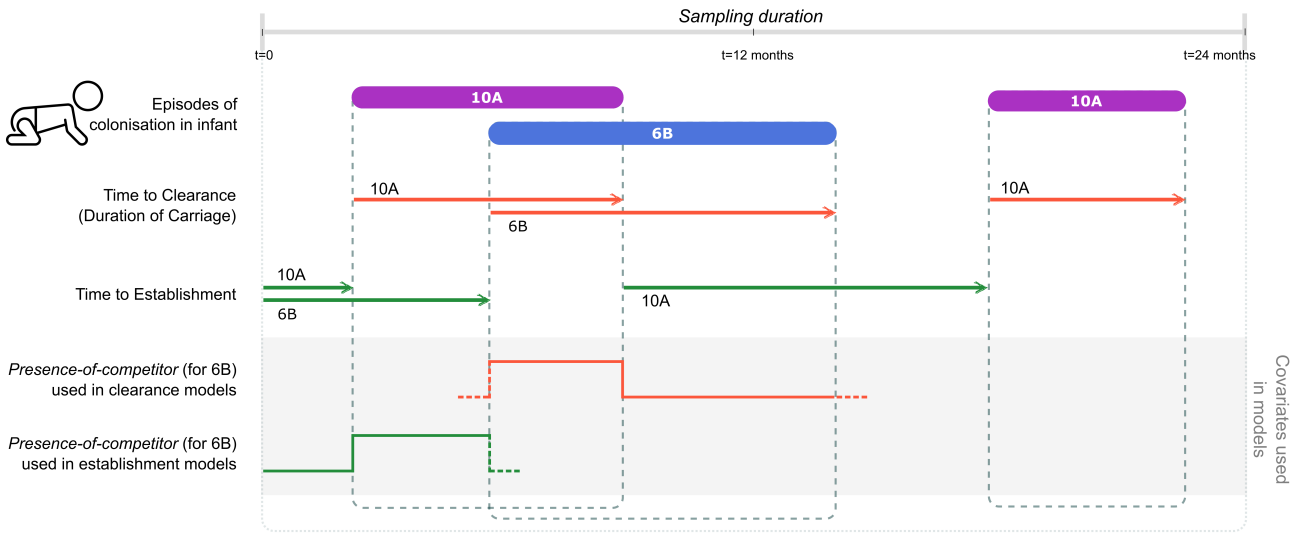

Figure 2. Definitions of covariates used in the survival analyses: Monthly sampling data of infants is used to identify the presence of serotypes over time in hosts, and infer the time to clearance and establishment of serotypes. Time to clearance is the duration from establishment to clearance of a serotype. Time to establishment is the duration from last clearance of that serotype to establishment. *Presence-of-competitor* variable is used to indicate whether there exists a competing strain in the host either during the time to clearance (orange lines) or time to establishment (green lines).

vaccine-targetted serotypes drives the increase in prevalence of non-targetted serotypes, with major consequences for the impact of the vaccine on e.g. disease prevalence (26–28). Competition between antibiotic-susceptible and resistant strains is important because it determines the behaviour of resistance frequencies (16, 29, 30).

In the case of antibiotic resistance, the ecology of strain competition is complex. The success of a resistant strain in a host depends on the fitness cost of resistance in the absence of antibiotics and the fitness benefit under antibiotic pressure. These effects can impact different components of strain fitness. At the epidemiological scale, fitness costs could manifest as reduced transmission rate, decreased duration of carriage, and in co-colonised hosts, as a decreased ability to compete with other strains. Numerous studies have estimated the fitness cost of resistance in experimental settings (e.g. (31)), typically in terms of effects on the growth rate. However, it is unclear how these estimates translate to humans *in vivo*, or how a reduction in growth rate affects epidemiological components of strain fitness. While fitness costs at the epidemiological scale can be inferred by fitting transmission models to surveillance and genomic data (32–35), these estimates depend on specific model structures. Untangling the effects of antibiotic resistance on various aspects of strain fitness in absence and presence of antibiotics remains a challenge.

To address these questions, we use a rich longitudinal dataset of carriage of *S. pneumoniae* in infants (12, 36, 37) to quantify the impact of within-host competition on various aspects of pneumococcal dynamics. The dataset has presence-absence information on pneumococcal serotypes carried in 737 infants sampled monthly for up to 2 years, along with additional information on antibiotic consumption in the infants and resistance phenotypes to seven antibiotics for a subset of all samples. We use survival analyses to quantify the effects of the presence of within-host competitors on two important epidemiological processes - clearance rate and establishment rate of serotypes, thereby linking the two scales of competition. Using the data on resistance status of strains, we quantify the within-host fitness cost and benefit of resistance. Our results advance our understanding of within-host competition between bacterial strains and provide a basis for improved prediction of pneumococcal – and in particular antibiotic resistance – dynamics.

Results

Survival models can explain pneumococcal carriage dynamics

We use data on pneumococcal carriage in 737 infants from (12), with nasopharyngeal swabs taken roughly once a month to identify the serotypes that they carried. Assuming complete accuracy of serotype detection, we infer 5529 episodes of colonisation by 71 different serotypes across all infants (see Figure S1 for a visualization of the data). We use these inferred colonisation episodes to quantify the effects of within-host processes on two epidemiological parameters in our analyses: clearance rate and establishment rate. Clearance rate is defined as the inverse of the average time between establishment of a strain and its clearance (i.e. the inverse of the duration of carriage). The establishment rate of a strain is defined as the inverse of the average time to establishment of the strain since the previous clearance of that strain in the host (Figure 2). This rate is related to, but not the same as, transmission rate: establishment rate is the product of the transmission rate and the density of transmitting hosts.

We use Cox Proportional Hazards modelling to quantify the effects of different strain-associated factors on clearance and establishment rates in independent models. Briefly, the Cox model describes the time to an event T (time to clearance, or time to establishment) using a hazard function as follows:

$$H(t|X) = H_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p) \quad \textbf{(1)}$$

Here, H denotes the hazard of clearance or establishment, H_0 denotes the baseline hazard function, X_p denotes the value of covariate p, and β_p represents the strength of covariate p in determining the hazard of the event. The function H_0 is not estimated by the Cox model.

Diversity in clearance rates across serotypes. First, we introduce age of infant at colonisation and serotype of the colonising pneumococci in a model for time to clearance. The different serotypes are included as covariates in the model and age is included as a stratifying variable (here, and in all subsequent analyses, see Methods). Stratification by age leads to an age-dependent baseline hazard function in the model. The hazard coefficients corresponding to the serotypes describe the inherent durations of carriage of the different serotypes. Serotypes with a lower hazard coefficient are those with a longer duration of carriage – i.e. a lower clearance rate. The results of this analysis (summarized in Figure S2 A,B) capture known patterns of pneumococcal ecology. For example, serotypes 6B, 19F and 23F, with the lowest hazard coefficients in the analysis, are also known to be long-colonisers from previous studies (38, 39).

Antibiotic consumption increases clearance rates. Next, we introduce information on antibiotic use in infants into the above model (see Methods for details about the data). We expect to observe faster clearance of serotypes in hosts due to antibiotic consumption. We define a time-dependent binary variable *drug-use* that tracks whether an infant consumes any antibiotic during each episode of colonisation. Here, the start dates of *drug-use* are taken from our dataset and the length of all treatments are assumed to be 7 days based on WHO guidelines on treatment of children (40) (sensitivity to the assumption discussed in Supplementary Note 6). We include *drug-use* as a time-varying covariate, serotypes as constant covariates, and age as a stratifying variable in the model. To enable model convergence, we removed serotypes with zero variance in survival outcomes (these are serotypes with low prevalence) from the analysis (see Methods). The results (Figure S2C) reveal a positive hazard coefficient of 0.48 (95% confidence intervals: CI=0.37 to 0.59) corresponding to *drug-use*. The exponential of the hazard coefficient gives the hazard ratio for the covariate and describes the change in hazard of the event under a unit change in the covariate. Thus antibiotic consumption is associated with an increase in the rate of clearance of strains by a factor exp(0.48) = 1.62, or 62% (CI=45% to 80%).

Within-host competition increases clearance rate of *S. pneumoniae*

We then look at the effect of within-host competition on the clearance rate of a serotype. To do this, we define a time-dependent variable *presence-of-competitor*, which tracks how many serotypes are carried in a host. *Presence-of-competitor* corresponding to a focal serotype at a given time equals 1 if there exists another serotype in the same host at that time, thus indicating whether the focal serotype experiences within-host competition (Figure 2). This is a binary variable: we assume that the focal serotype experiences a constant competitive pressure independent of the number of co-colonising serotypes or their abundances. We perform time-varying Cox proportional hazards regression using *presence-of-competitor* as the time-varying covariate along with serotype of the focal strain, while correcting for infant age by stratification (Figure 3A). We use these covariates since the model that includes serotype has a lower partial-AIC compared to one that does not (details in Supplementary Note 7). The hazard coefficient corresponding to *presence-of-competitor* is 0.28 (CI=0.21 to 0.35), indicating an increase in clearance rate in the presence of another serotype (Figure 3A). Thus, co-colonisation is associated with an increased risk of clearance of exp(0.28) = 1.33, or 33% (CI=24% to 42%).

Next, we looked further into what determines the outcomes of within-host competition. To test whether serotypes differ in their within-host competitive abilities, we introduce interaction terms between each serotype and *presence-of-competitor*. These interaction terms capture the extent to which the clearance rate of each focal serotype is affected by the presence of another competing serotype. We perform a time-varying Cox proportional hazards model with the following covariates: *presence-of-competitor*, serotype of the focal strain, and all of the interaction terms. To determine if the intrinsic clearance rate (i.e. clearance rate in the absence of competitors) of a serotype predicts the effect of a co-colonising competitor, we fit a linear model (using orthogonal distance regression that accounts for uncertainty in both variables – see Methods) between the hazard coefficients of serotypes and their interaction terms. We find a negative relationship (slope=-0.40, CI=-0.60 to -0.20), indicating a negative relationship between inherent rate of clearance and the change in clearance rate under within-host competition (Figure 3B). These results suggest that faster clearing strains may be better at within-host competition than longer-carried ones, indicating a potential trade-off between intrinsic duration of carriage and within-host competitive ability. Overall, the above analysis indicates that serotype-specific competitive abilities can partially predict the outcome of within-host competition.

We also tested whether order of colonisation of strains in the host can affect clearance under competition. We refer to this as a priority effect, in which a resident strain can have a competitive advantage in clearance over a newly established strain. We replace *presence-of-competitor* with two covariates that encompass information on order of establishment of the serotypes: *presence-of-resident-competitor* and *presence-of-invader-competitor*. *Presence-of-invader-competitor* indicates that the focal strain is the resident serotype and conversely, *presence-of-resident-competitor* indicates the focal serotype is the invader. A Cox proportional hazards model with serotype of the focal strain, *presence-of-resident-competitor* and *presence-of-invader-competitor* as covariates reveals that the presence of a resident serotype increases the risk of clearance of the competitor significantly more than an invader serotype (Figure 3C). This analysis is potentially confounded because acquisition of an additional serotype is more likely to occur during longer episodes of carriage – long episodes of carriage are therefore more likely to be resident than short ones. To correct for this effect, we compare the coefficients from our analysis to their expected null distributions when episodes of colonisation in infants are randomly reshuffled (histograms in Figure 3C). This reshuffling allows us to generate the null distributions of the hazard coefficients in absence of competitive effects, i.e. where effects are only due to the confounding

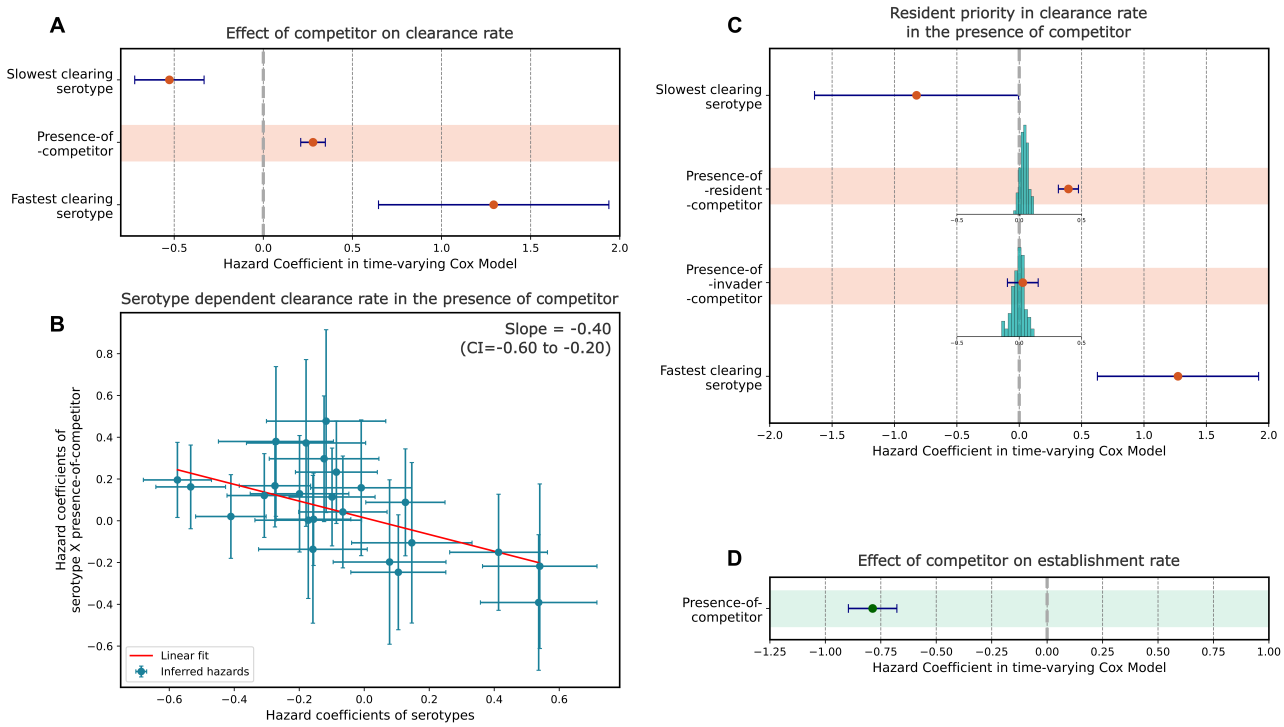

Figure 3. Models that show the effects of within-host competition. A-C correspond to models for clearance rate, and D corresponds to a model for establishment rate. Error bars show 95% confidence intervals. (A) Does the presence of a competitor affect the clearance rate of a serotype? *Presence-of-competitor* has a positive hazard, indicating increased clearance rate in the presence of competitors. (B) Does the increase in clearance in the presence of competitors depend on the serotype of the focal strain? The plot shows a linear model that maps hazard coefficients of serotypes to hazard coefficients of interaction terms between serotype and *presence-of-competitor*. Each blue dot corresponds to one serotype. Error bars in this plot are standard errors. The red line is the fit to a linear model. (C) Do resident strains have a competitive advantage for clearance? *Presence-of-resident-competitor* is associated with an increased clearance rate, while *presence-of-invader-competitor* is not. The histograms represent the null distribution of the variables when episodes of colonisation occur randomly within infants. (D) Does the presence of a competitor affect establishment rate of a strain? *Presence-of-competitor* has a negative hazard on establishment, indicating a reduced rate of establishment into an occupied host niche.

described above. The coefficient of *presence-of-resident-competitor* (0.40, CI=0.32 to 0.48) is significantly higher than that expected from its null distribution (p-value<0.01). The presence of an invader strain is not associated with an increase in clearance rate. This analysis suggests that pneumococcal within-host competition is asymmetric, with the invader strain incurring a higher clearance risk compared to the resident strain. This result is qualitatively sensitive to how the start and end of a colonisation episode are defined (more in Supplementary Note 2).

Within-host competition reduces establishment rate of *S. pneumoniae*

Next, we look at the effect of within-host competition on the rate of establishment of a serotype. The establishment rate of a serotype is defined in terms of the time since the last clearance event of that serotype from the host (Figure 2). We test whether the establishment rate is modified by the presence of a resident competitor in the host. Here, we are interested in the presence of potential competitors in the host during the 'time to establishment' of the serotype. We define a time-dependent variable named *presence-of-competitor* that varies along the time to establishment of a serotype, and tracks the status of the host into which it eventually establishes: in the time interval preceding establishment of the focal serotype, *presence-of-competitor* equals one when the host is carrying at least one competing strain (i.e. non-focal serotype). As in the previous analyses, we run a time-varying Cox proportional hazards model on the data using *presence-of-competitor* as a time-varying covariate. We do not include the serotype of the incoming strain as a covariate in this analysis since serotype-differences in establishment rates depend on the prevalence of the serotypes in the population. Instead, we add serotype along with age of infant as a stratification variable that allows for a different baseline hazard function for each serotype. This adjusts for the affect of serotype without estimating its effect on the outcome.

This analysis (Figure 3D) reveals a significant decrease in establishment rate in the presence of additional serotypes in the host. The *presence-of-competitor* covariate has a negative hazard coefficient of -0.79 (CI =-0.90 to -0.68) and a hazard ratio of 0.46 (CI=0.41 to 0.51). This corresponds to establishment into a host occupied by a competitor occurring at about half the rate (reduction by 1-0.46 = 54%, CI= 49% to 59%) compared to establishment into an empty host niche.

Alhough the effect of serotype of the incoming focal strain could not be estimated, we can test whether the serotype of the resident competitor affects the establishment rate of the focal strain. We run a time-varying Cox proportional hazards model using *presence-of-competitor* and serotype of the resident strain as covariates, while adjusting for age of host and serotype of the focal strain by including them as stratification variables. The model shows variability in how well resident serotypes prevent colonisation by a new serotype (Figure S4). Serotypes with a negative hazard

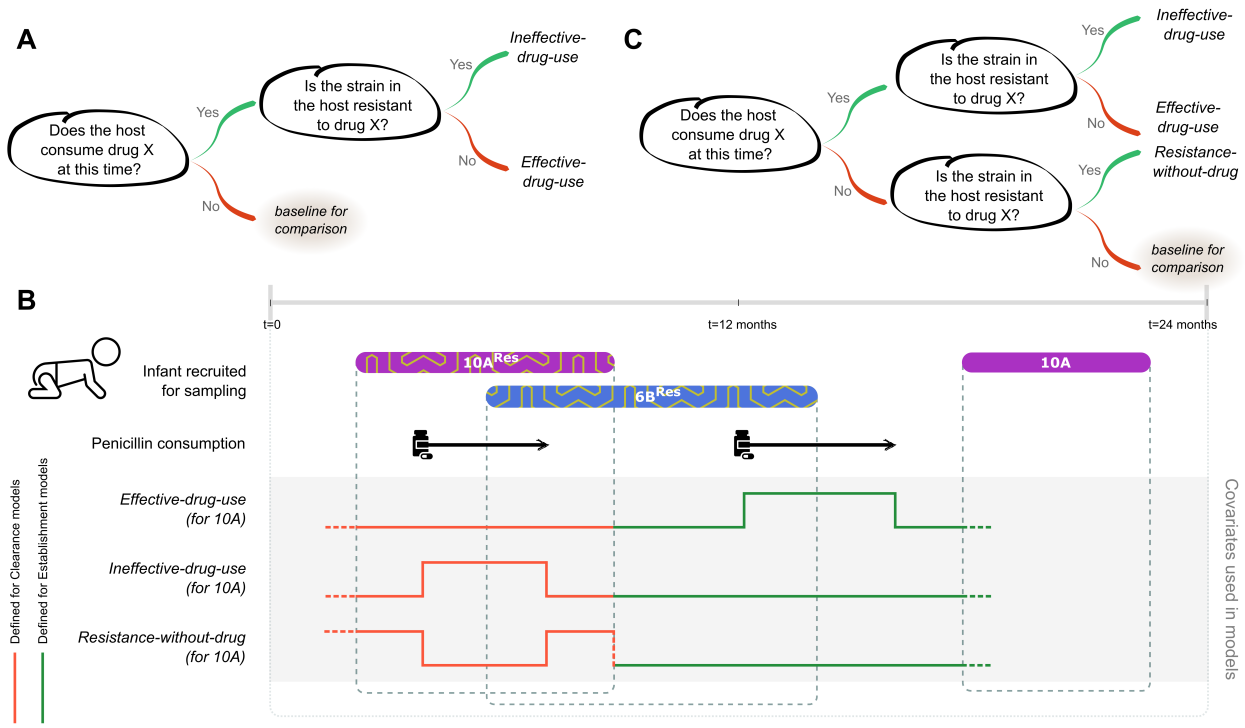

Figure 4. Definitions of covariates relating to drug usage and antibiotic resistance: (A,C) Schematics explaining how the categorical variables *effective-drug-use*, *ineffective-drug-use* and *resistance-without-drug* are defined based on the set of drugs consumed by the host and resistances carried on strains. Since these three variables are defined analogous to levels of a categorical variable, their effect sizes are to be interpreted relative to the fitness effect corresponding to '*baseline for comparison*'. (B) An example showing how the values of the three covariates vary with time. The first two episodes of carriage are resistant to penicillin. The orange lines indicate the value of the covariates for the first 10A serotype as used in clearance models. The green lines indicate the value for the second 10A serotype as used in establishment models.

coefficient may be better at competition during colonisation and preventing establishment of incoming serotypes, while those with a positive coefficient may be worse at preventing establishment. A likelihood ratio test shows an overall significant effect of serotype of competitor on establishment rate (*p-value:* 1.2×10^{-13}). In summary, the serotype of within-host competitors can affect the rate of establishment of incoming pneumococcal strains.

Antibiotics and antibiotic resistance affect clearance rates of strains

Next, we shift our focus to understanding how antibiotic consumption and antibiotic resistance affect pneumococcal dynamics. Antibiotic consumption is associated with an overall 62% increase in clearance rate *(Section: Survival models can explain pneumococcal carriage dynamics)*, but this effect may depend on the resistance profiles of these strains. Conversely, the effect of resistance on clearance and establishment can depend on antibiotic consumption: resistance is potentially beneficial under antibiotic exposure, but costly in the absence of antibiotics. We seek to separately quantify the effects of drug consumption and resistance on clearance rates in our dataset (effects on establishment discussed in Supplementary Note 3). In this section, we focus on the effects of antibiotic consumption on the dynamics of antibiotic-resistant versus susceptible strains.

The dataset contains information on 13 different drug treatments that the infants are exposed to, but resistance status of the strains to only seven antibiotics. We map the drug treatments to the corresponding resistance phenotypes based on antibiotic class (see Methods), and define a categorical variable for drug use that takes into account both the resistance status of the focal strain and antibiotic consumption in the host (Figure 4A). If a host consumes antibiotics that the strain is resistant to, we refer to the drug usage as *ineffective-drug-use*; the variable measures the effect of antibiotics on a resistant strain. If the host consumes antibiotics that the strain is susceptible to, we refer to it as *effective-drug-use*; the variable measures the effect of antibiotics on a susceptible strain. These variables quantify the effects of drug usage on strain fitness relative to the 'baseline' condition (Figure 4A) corresponding to fitness of strains that experience no drug pressure. Both the variables are defined with respect to the focal strain (Figure 4B).

To analyse the effects of antibiotic consumption on clearance rate, we use serotype, *presence-of-competitor*, *effective-drug-use*, and *ineffective-drug-use* as covariates and run a time-varying Cox model on time to clearance data (these covariates produce the lowest AIC compared to a subset of these covariates - See 7). The hazard coefficient of *effective-drug-use* on clearance is 0.53 (CI=0.41 to 0.66) (Figure 5A). This shows that consumption of drugs is associated with a 71% (CI=51% to 93%) increase in the clearance rate of susceptible strains compared to no drug consumption. The hazard coefficient of *ineffective-drug-use* is also positive at 0.34 (CI=0.14 to 0.53, i.e. a 40% increase, CI=15% to 70%), suggesting that drug consumption increases clearance even for strains resistant to that drug. However, the hazard of *effective-drug-use* on clearance is higher that that of *ineffective-drug-use*, suggesting resistance does reduce drug-induced clearance (the size of

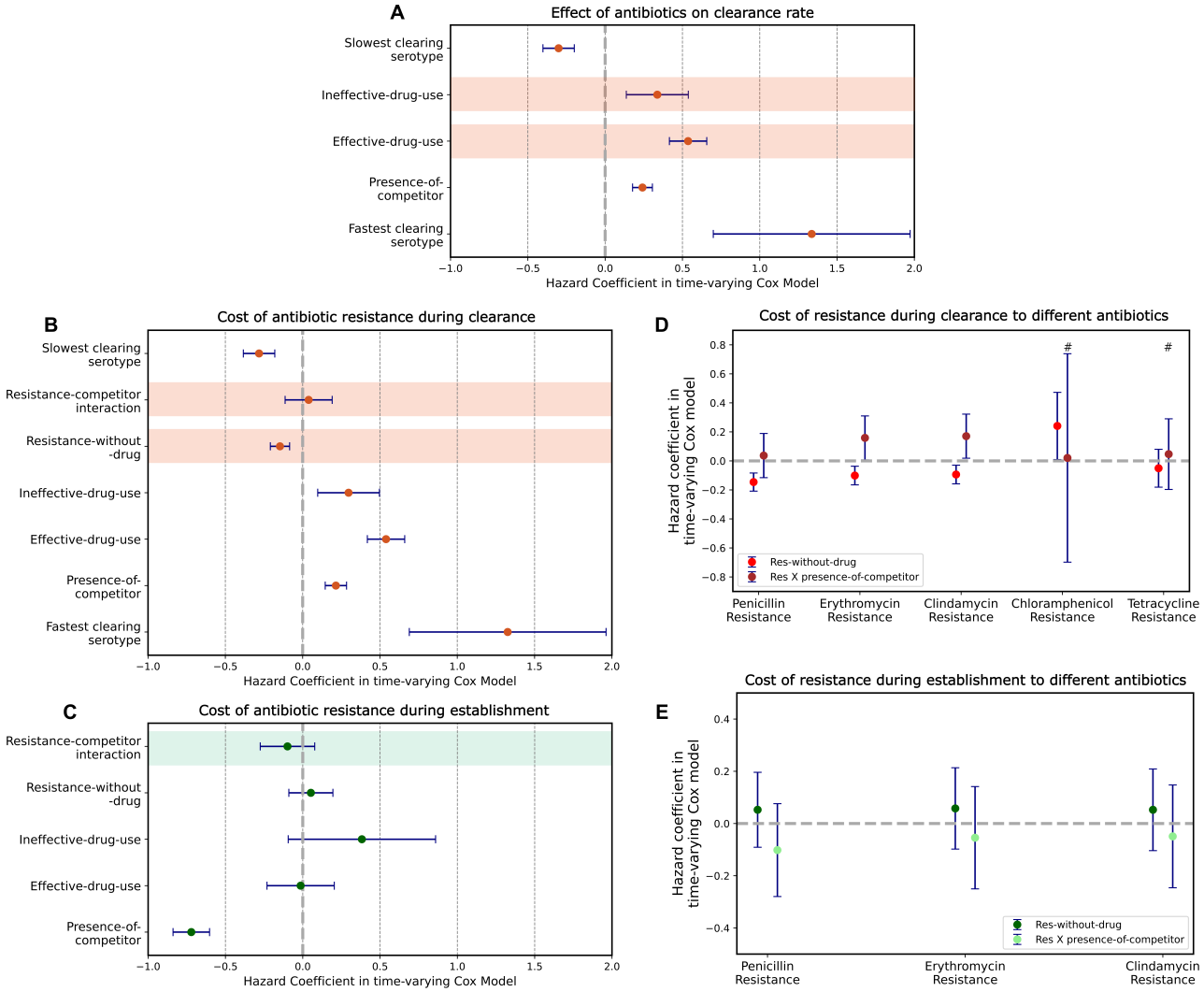

Figure 5. Models that show the effects of antibiotics and resistance on pneumococcal dynamics. A,B,D are models for clearance rate, while C,E show models for establishment rate. Error bars show 95% confidence intervals. (A) Does antibiotic consumption affect the clearance rate of resistant and susceptible strains? Here, *effective-drug-use* and *ineffective-drug-use* have positive hazards on clearance, indicating that antibiotics increase the rate of clearance of both susceptible and resistant strains. (B) Does resistance have a cost on clearance or within-host competitiveness? *Resistance-without-drug*, which corresponds to resistance on strains that are not exposed to antibiotics, has a negative clearance hazard, indicating reduced clearance rates. The interaction term between *resistance-without-drug* and *presence-of-competitor* which captures the competitive cost of resistance has a positive non-significant hazard coefficient. (C) Does resistance have a cost on within-host competition during establishment? The interaction term here captures the competitive cost of being resistant rather than susceptible when establishing into an occupied host niche. (D) The analysis in C is repeated for different resistances. # corresponds to effect sizes from models that do not include *effective-drug-use* and *ineffective-drug-use* since these resistances have no corresponding drug treatment in the dataset. See more in Supplementary Note 5. Only those resistances with enough data for model convergence are shown. (E) The analysis in C is repeated for resistance to different antibiotics.

this reduction is sensitive to our assumption about the length of all antibiotic treatments being 7 days, and is significant for 10 or more days; see more in Supplementary Note 6). Overall, the above analyses show that antibiotics clear both susceptible strains and resistant strains, but that resistance is associated with a reduced effect.

The effects of antibiotic use on clearance rate depend on the specific antibiotic consumed. In Supplementary Note 4, we show the results of the above analyses stratified by antibiotic. For amoxicillin/ampicillin, the hazard coefficients of *effective-drug-use* (0.55, CI=0.41 to 0.68) and *ineffective-drug-use* (0.34, CI=0.13 to 0.54) are positive. This is also the case for all other antibiotics, but the confidence intervals overlap zero (Supplementary Figure S5B). Amoxicillin and ampicillin, which make up most of the consumption in the dataset (2024 prescriptions during 2441 episodes of treatment), likely drive the overall effect of drugs we see.

Costs of resistance in the absence of antibiotics

Resistance to antibiotics is potentially associated with a fitness cost, which could affect various components of fitness, including the transmission rate, clearance rate and/or within-host competitiveness. Here, we quantify the effects of resistance on clearance rate and within-host competitiveness in the absence of antibiotics (see Supplementary Note 5 for antibiotic-use-blind analyses of the effects of resistance). We cannot make inferences about transmission rate, as observed establishment rates are dependent on the overall prevalence of resistant and susceptible strains in the population, for which we do not have reliable data.

We expand the categorical variables used in the preceding section to include a new variable for antibiotic resistance of the focal strain (see Figure 4C) called *resistance-without-drug*. Note that the 'baseline' scenario in Figure 4C is dif-

ferent from that used earlier in 4A. *Resistance-without-drug* tracks both the treatment status of the host over time and its relation to the resistance of the focal strain (Figure 4B).

We first ask whether resistance imposes a cost on clearance rate, i.e., whether resistant strains are cleared faster than susceptible strains in the absence of antibiotics. The inference of such a cost, however, is complicated by an epistatic interaction between resistance and duration of carriage, which leads to resistance being more common on long-carried strains (41). The ability to detect a cost on clearance rate hinges on the relative strength of this cost compared to the association of resistance with low clearance rates.

To measure the effect of resistance on clearance rates, we use a time-varying Cox proportional hazards model with the following explanatory variables: serotype of focal strain, *presence-of-competitor*, *effective-drug-use*, *ineffective-drug-use* and *resistance-without-drug* (we use this model since these covariates produce a lower AIC than models with a subset of these covariates - see Supplementary Note 7). The hazard of *resistance-without-drug* here indicates the effect of a resistant strain on clearance rate compared to a susceptible strain in the absence of antibiotics. The hazard coefficient of *resistance-without-drug* is -0.14 (CI=-0.20 to -0.08, 13% lower rate than susceptible strains (CI=8% to 18%))(Figure S8, also Figure 5B), indicating that resistance is associated with a lower clearance rate, i.e. a longer duration of carriage. Thus, any potential cost of resistance here is fully masked by the above-mentioned association. We additionally perform this analysis for each resistance separately to measure costs of resistance to different antibiotics. We find similar negative coefficients for resistance to penicillin, erythromycin and clindamycin (Figure 5D). However, resistance to chloramphenicol is associated with a positive coefficient of 0.24 (CI= 0.01 to 0.47) corresponding to a 27% (CI= 1% to 60%) higher clearance rate of chloramphenicol resistant strains compared to susceptible strains.

Next, we analyse the effect of antibiotic resistance on the competitive ability of co-colonised strains. This competitive cost of resistance can manifest as an increased effect of the presence of a within-host competitor on either the clearance rate or the establishment rate of a resistant strain (compared to a susceptible strain).

To test whether resistance is associated with an increased effect of competitors on clearance of the focal strain, we use a time-varying Cox model with the following covariates: serotype of focal strain, *presence-of-competitor*, *resistance-without-drug*, *ineffective-drug-use*, *effective-drug-use*, and additionally an interaction term between *resistance-without-drug* and *presence-of-competitor*. This interaction term captures how the hazard of clearance of an antibiotic resistant strain in an antibiotic-free host changes due to the presence of a competing strain, thereby reflecting the competitive cost of resistance. The interaction term has a hazard coefficient of 0.04 (CI=-0.11 to 0.19) (Figure 5B): there is an overall 4% (CI= -10% to 20%) change in clearance rate of resistant strains under within-host competition compared to susceptible strains. We also repeat this analysis for resistance to each antibiotic separately where possible (Figure 5D). Here, we find some evidence for a cost of resistance on competitive ability during clearance. Resistance to erythromycin and clindamycin is associated with an increase in the clearance rate of a strain in the presence of competitors by 17% (CI=1% to 36%, coefficient=0.16, CI=0.01 to 0.31) and 19% (CI=2% to 38%, coefficient=0.17, CI=0.02 to 0.32) respectively.

Finally, we test whether resistance is associated with an increased effect of competitors on the establishment of the focal strain, using a time-varying Cox model that includes *presence-of-competitor*, *resistance-without-drug*, *ineffective-drug-use*, *effective-drug-use*, and an interaction term between *resistance-without-drug* and *presence-of-competitor* as covariates. The interaction term here has a hazard coefficient of -0.10 (CI=-0.27 to 0.08) (Figure 5C), corresponding to a 10% (CI=-8% to 24%) decrease in establishment rate of resistant strains into occupied host niches compared to susceptible strains. The analysis for resistance to different antibiotics separately is shown in Figure 5E. Here again, the interaction terms are negative and non-significant for all resistances tested. Overall, we do not see a clear evidence for cost of resistance on competitive ability during establishment.

Discussion

Competition affects clearance and establishment. We analysed the largest-to-date longitudinal dataset on pneumococcal carriage in infants to quantify the effects of the presence of within-host competitors on the clearance rate (or the duration of carriage) and establishment rate of serotypes. The presence of another serotype in the host was associated with a 33% (CI=24% to 42%) higher rate of clearance and a 54% (CI=49% to 59%) lower rate of establishment by new serotypes. Our results support the hypothesis that serotypes compete with each other to persist within a host. The competitive effect observed on establishment rate is in line with previous studies which have reported reduction in the rates of co-colonisation of strains compared to rates of initial colonisation (17, 42). These estimates are derived from fitting longitudinal data to mathematical models and range from around 50% to 90% reduction in establishment rates depending on the dataset used, serotypes tested and assumptions made in the model. The competition during establishment is analogous to the term *colonisation resistance* (43, 44) mentioned in other contexts, which refers to how resident bacteria reduce the likelihood of colonisation by new bacterial species or strains.

The effect of competition on clearance, on the other hand, has not been extensively discussed in previous work on pneumococcal dynamics. Few studies (with much smaller sample sizes) reported no significant effects of the presence of competitors on the clearance of serotypes (45, 46). Auranen et. al. (46) analysed an epidemiological dataset similar to this study using a multi-state Markov model. They reported a 10-times reduction in establishment rate due to within-host competition, but clearance rate decreased by a factor of 0.81 (CI=0.48 to 1.26) due to competition. While mathematical models commonly account for a lower rate of co-colonisation using an efficiency factor on transmission rate, they often assume constant clearance rates independent of the presence of co-colonisers (15–18). Our work emphasizes the significant hazard of within-host competition on accelerating clearance and its potential importance

in pneumococcal dynamics.

Serotype hierarchies during within-host competition We find that within-host competitive effects can be predicted by the identity of the competing serotypes. Biochemical properties of the polysaccharide capsule that determine a serotype are known to affect its fitness (47). Longer-carried serotypes have been previously found to prevent co-colonisation better than short-carried serotypes (42). Our study finds that a similar correlation also exists for competition at clearance, but in the opposite direction: serotypes with longer durations of carriage are associated with a larger reduction in duration of carriage in the presence of another competing serotype. This is surprising since longer serotypes have thicker capsules that increase other aspects of their fitness (47). Our result suggests that there may exist different mechanisms of between-serotype competition during establishment and during clearance.

The association between longer inherent duration of carriage and faster clearance by competitors suggests a potential trade-off between fitness in co-colonised hosts and fitness in singly colonised hosts, which could partially explain the diversity of serotypes observed in *S. pneumoniae*. The benefit gained from having an advantage during co-colonisation depends on how often a strain experiences co-colonisation, and thus increases with the prevalence of competing strains. As a result, this type of trade-off gives rise to negative frequency-dependent selection and thus contribute to the maintenance of diversity (48). Within-host competition has been linked to strain-diversity in other species as well (49, 50), through different mechanisms. For example, a recent longitudinal analysis of *E. coli* colonisation (50) also finds that the presence of within-host competitors increases clearance rate, and that this effect is greater when the competitors are from the same phylogroup. Stronger within- than between-phylogroup competition would thus contribute to maintaining phylogroup diversity.

Additionally, we find that within-host competitive outcomes depend on the order of colonisation of serotypes in the host: during co-colonisation, resident strains have a lower risk of clearance than invader strains. A recent experimental study showed such priority effects in competition during establishment in *S. pneumoniae* (51): when inoculated with multiple serotypes, mice carried higher densities of serotypes that were given an initial head-start in colonisation. This may also translate into an advantage in clearance for the early-coloniser if within-host densities post establishment affect duration of carriage. Our analysis provides direct evidence for such a priority effect on clearance rate.

Antibiotics affect clearance rates We show that consumption of amoxicillin or ampicillin is associated with a 73% (CI=51% to 97%) increase in the clearance rate of susceptible strains. Interestingly, we find that consumption of these drugs is also associated with increased clearance of resistant strains (by 45%, CI=15% to 69%). This might be because antibiotics can reduce the replication rate of resistant strains, even if the antibiotic concentration is not as high as the minimum inhibitory concentration (based on which resistance is binarized in our analyses). The slower replication in conjunction with other mechanisms (for instance, host immunity or interspecies competition) could accelerate the clearance of resistant strains. The difference in antibiotic-induced clearance between susceptible and resistant strains (by 28 percentage points for amoxicillin/ampicillin) quantifies the benefit of antibiotic resistance. Although we see that resistant strains are cleared slower than susceptible strains by antibiotics, the evidence for this fitness benefit is not strong due to the large confidence intervals. Infants in our data also consume multiple drugs at the same time (19% of treatment episodes), but our methodology is blind to the differential effects of combinations of drugs on strain dynamics, which been described recently elsewhere (52).

Fitness costs of resistance We estimated fitness costs of antibiotic resistance in the absence of antibiotics on (1) the clearance rate of strains, and (2) their competitive ability during clearance and establishment. Overall, the evidence we find is mixed.

The detection of any clearance-associated fitness costs is confounded by the previously described association between antibiotic resistance and long duration of carriage (41, 53). This association is thought to arise because resistance is more beneficial for longer-carried strains (41, 54), rather than due to the effects of resistance on the duration of colonisation. Indeed, our results replicate this finding and show, unlike previous analyses, that the association holds even when focusing on untreated hosts. The exception is chloramphenicol resistance, for which we see an association between resistance and shorter duration of colonisation. This could arise due to a large clearance-associated fitness cost for chloramphenicol. However, as we find this effect for only a single antibiotic, it should be interpreted with caution.

We find some evidence for a cost of resistance on within-host competitiveness. Specifically, we find that the presence of a within-host competitor is associated with a greater effect on the clearance rate of erythromycin (17%, CI=1% to 36%) and clindamycin (19%, CI=2% to 38%) resistant strains compared to susceptible strains. For the other resistances, as well as for the overall estimate across all resistances, the direction of the effect is consistent with a fitness cost, but the confidence intervals overlap zero. These results qualitatively align with the estimates of within-host competitive advantage of susceptible strains inferred in a previous study, which also found much higher costs for macrolides than penicillins (14).

The difficulty in identifying and quantifying the fitness cost of resistance is intriguing. Resistance frequencies do not appear to be increasing towards fixation (i.e. 100%) (16). Indeed, a systematic analysis of resistance trajectories in Europe over two decades shows evidence of resistance frequencies stabilising at intermediate values across a range of bug-drug combinations, and no evidence of a wide-spread increasing trend (55). Such trends strongly suggest resistance is associated with a fitness cost; otherwise, we would expect to observe increase towards fixation. It is unclear why this is not apparent in this study. One possible explanation is that the cost of resistance is primarily associated with transmission. To detect effects on transmission rate from this dataset, we would need reliable data on the overall resistance prevalence in the study setting. Another possibility is that because fitness cost is a complex and heterogeneous trait, which depends on both environment and genetic background, its magnitude is very difficult to quantify. This is consistent with evidence from other studies on fitness cost,

including experimental studies, often being mixed (56).

Conclusion Using longitudinal data on *S. pneumoniae* carriage in infants, we performed detailed analysis of the effects of within-host competition and antibiotic resistance on pneumococcal epidemiology. The key limitation of our work is that, as with all observational studies, we detect associations, not causation. Nevertheless, when considered in the context of previous work on bacterial ecology, the effects we find have very plausible causal explanations. Overall, our study provides insights into the role of within-host competition on pneumococcal epidemiology. Such insights play an important role in progressing towards a better understanding of pneumococcal and antibiotic resistance dynamics.

Methods

The dataset The Maela Pneumococcal Collaboration collected nasopharyngeal samples from 952 mother-infant pairs from the Maela refugee camp at the Thai-Myanmar border roughly once a month between 2007 and 2010. Excluding dropouts and stillbirths, samples were collected from 737 infants during their first two years of life. The samples were serotyped according to the latex sweep method or WHO pneumococcal carriage detection protocols (234 of the initially recruited infants were serotyped by WHO culture and 721 by latex sweeps). 71 different serotypes were identified in the study, and pneumococci lacking cell wall capsules were tagged as Non-typeable (NT). The metadata also included episodes of illness, hospital visits and antibiotic consumption. A total of 2441 episodes of illness were recorded and treatment corresponded to one of the following: amox/ampicillin, ciprofloxacin, gentamicin, erythromycin, cloxacillin, ceftriaxone, clindamycin, cefotaxime, TB(tuberculosis) treatment, cotrimoxazole, cefixime, penicillin, nitrofurantoin. Additionally, a subset of samples were chosen for the detection of phenotypic resistance to seven antibiotics. 33.7% of all samples have been tested for resistance to the following seven antibiotics: ceftriaxone, cotrimoxazole, clindamycin, erythromycin, penicillin, tetracycline, and chloramphenicol. Resistance to at least one antibiotic was detected in 30.4% of all samples.

Data Pre-processing We include only data of infants in all our analysis. We assume serotype detection to be fully accurate to determine duration of carriage of serotypes. When a serotype is detected in a sample at time T_1 when it is not detected in the preceding sample at time T_0, then the serotype is assumed to have established at time $0.5x(T_0+T_1)$. If a serotype is already detected in the earliest sample collected from the infant, then the establishment is unobserved (censored). Similarly, a serotype detected continuously until sample T_2 but not in the succeeding sample at T_3 is assumed to be cleared at time $0.5x(T_2+T_3)$ (Sensitivity of results to this assumption is discussed in Supplementary Note 2). Some monthly swabs were missing in the data, but these do not alter the above estimation of establishment and clearance. The duration between an establishment and clearance is the *Time to Clearance* of a strain. To get the *Time to Establishment*, we consider the preceding time of clearance of that serotype until the time of establishment of the focal strain. This is the time period during which the serotype could potentially establish in the host.

To infer the resistance status of the different episodes of carriage in our data, we assume a strain to be antibiotic-resistant if it was labelled as 'INTERMEDIATE' or 'RESISTANT' in the data, and antibiotic-susceptible if labelled as 'SUSCEPTIBLE'. We also assume that resistance is not lost during an episode of carriage, i.e., any conflicts in resistance status of an episode of carriage are resolved by treating the entire episode as resistant. In this way, we were able to infer the resistance status of 1944 out of the 5529 episodes of colonisation in all infants.

To include both drug use and resistance data in our analysis, we create a map between the two based on antibiotic class. *Drug(in italics)*: resistance status =

amoxicillin/ampicillin: penicillin,

erythromycin: erythromycin,

cloxacillin: penicillin,

ceftriaxone: ceftriaxone,

clindamycin: clindamycin,

cefotaxime: ceftriaxone,

cotrimoxazole: cotrimoxazole,

cefixime: ceftriaxone,

penicillin: penicillin

Resistance profiles against the following drug treatments are not known: ciprofloxacin, gentamicin, TB treatment and nitrofurantoin. An X:Y link in the above map is used to signify that when the host carries Y-resistant strains, then consumption of drug X may not affect those strains. For the four drug treatments not included in the above map, we assume there is no resistance.

Survival Analyses We use survival analysis to model our data since some establishment and clearance events of strains are censored. All survival analyses are performed in python using the *lifelines* package (57). We use the Cox proportional hazards model since it is non-parametric and hence makes the least assumptions about the underlying baseline hazard function. Models of clearance include right censoring when clearance events are unobserved, while models of establishment include left censoring when start of establishment periods are unobserved.

Age of the infant at colonisation is added as a categorical stratification variable in all analyses. We do this since age does not follow the proportional hazards assumption to be included as covariates, i.e., its effect on clearance and establishment vary continuously with time. Age can belong to one of four categories: up to 6 months, 6 to 12 months, 12 to 18 months, and above 18 months.

In the analyses, where required, we add the 71 serotypes as 70 dummy variables (and one reference serotype 18C, randomly chosen). Thus all serotype coefficients capture hazards relative to serotype 18C. Non-typeable serotypes (NT) are not included in the analysis. Some serotypes can cause model convergence issues due to low total prevalence in the population (technically known as perfect separation). In these cases, these serotypes are removed from the focal serotypes before the analysis but they are included as po-

tential within-host competitors. In this way, 49 serotypes are included in the models explaining the effect of competitors, and 56 serotypes for models describing overall drug and resistance effects. For models that include any resistance-related covariates, we use only those episodes of colonisation whose resistance information is known to run the model. In all cases where an infant consumes multiple drugs, the model assumes that the drugs do not interact with each other and only affect its corresponding resistances as mentioned in the drug-resistance map above. All variables we include in our models as covariates are binary and unit-less.

We use the Akaike Information criterion (AIC) to determine which model to use for each analysis. Our approach is as follows: to estimate the effect of a given variable X, we use a model that includes X and any covariates that reduce the AIC when added to the model. We calculated the AIC in each of the models used in the main text, and show that models with only a subset of covariates have a higher AIC than the model used in a given analysis in Supplementary Note 7. Additionally, models with more variables than necessary also increase the AIC (see 7), and therefore are less parsimonious. In these calculations, serotype is taken to be one categorical variable, so that AIC for models with subsets of serotype dummy variables are not estimated.

To analyse the effect of serotype identity on clearance rate in the presence of competitors, we used a linear regression approach: we performed Orthogonal distance regression between the hazard coefficients of serotype and those of interaction terms between serotype and *presence-of-competitor*. This method differs from Simple Linear Regression by taking into account uncertainty in both dependent and independent variables (minimises errors in both directions). We are interested in whether the slope of this relationship is significantly different from zero.

The analysis of priority effects during clearance in the presence of competitors can be confounded by the fact that longer durations of carriage are more likely to establish first in the host. To correct for this effect, we generated a null distribution of hazard coefficients of *presence-of-resident-competitor* and *presence-of-invader-competitor* by reshuffling the data 100 times. Each time, the episodes of carriage within each infant are reshuffled randomly within the sampling duration. The hazard coefficients inferred from data are compared to the null distributions. The p-value is the fraction of values in the null distribution that is at least as extreme as the hazard coefficient from data.

Code and Data Availability The data on pneumococcal carriage in infants was obtained from this repository. The code corresponding to results is available here.

Acknowledgements

We thank Lucas Graz from the Seminar for Statistics (SfS) at the ETH Zurich for statistical advice and Christophe Fraser for discussions. This study was funded by the Swiss National Science Foundation (SNSF) project number PR00P3_201618.

Bibliography

1. T Ferenci, Trade-off mechanisms shaping the diversity of bacteria. *Trends in Microbiology* **24**, 209–223 (2016).
2. RA Welch, et al., Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences* **99**, 17020–17024 (2002).
3. NL Hiller, et al., Comparative Genomic Analyses of Seventeen *Streptococcus pneumoniae* Strains: Insights into the Pneumococcal Supragenome. *Journal of Bacteriology* **189**, 8186–8195 (2007).
4. J Friedman, LM Higgins, J Gore, Community structure follows simple assembly rules in microbial microcosms. *Nature Ecology & Evolution* **1**, 1–7 (2017) Publisher: Nature Publishing Group.
5. S Freilich, et al., Competitive and cooperative metabolic interactions in bacterial communities. *Nature Communications* **2**, 589 (2011) Publisher: Nature Publishing Group.
6. T Lazebnik, S Bunimovich-Mendrazitsky, Generic approach for mathematical model of multi-strain pandemics. *PLOS ONE* **17**, e0260683 (2022) Publisher: Public Library of Science.
7. R Kouyos, E Klein, B Grenfell, Hospital-Community Interactions Foster Coexistence between Methicillin-Resistant Strains of Staphylococcus aureus. *PLOS Pathogens* **9**, e1003134 (2013) Publisher: Public Library of Science.
8. MT Meehan, DG Cocks, JM Trauer, ES McBryde, Coupled, multi-strain epidemic models of mutating pathogens. *Mathematical Biosciences* **296**, 82–92 (2018).
9. X Wang, S Panchanathan, G Chowell, A Data-Driven Mathematical Model of CA-MRSA Transmission among Age Groups: Evaluating the Effect of Control Interventions. *PLOS Computational Biology* **9**, e1003328 (2013) Publisher: Public Library of Science.
10. M Aguiar, et al., Mathematical models for dengue fever epidemiology: A 10-year systematic review. *Physics of Life Reviews* **40**, 65–92 (2022).
11. JR Shak, JE Vidal, KP Klugman, Influence of bacterial interactions on pneumococcal colonization of the nasopharynx. *Trends in Microbiology* **21**, 129–135 (2013).
12. G Tonkin-Hill, et al., Pneumococcal within-host diversity during colonization, transmission and treatment. *Nature Microbiology* **7**, 1791–1804 (2022) Publisher: Nature Publishing Group.
13. T Khawaja, et al., Deep sequencing of Escherichia coli exposes colonisation diversity and impact of antibiotics in Punjab, Pakistan. *Nature Communications* **15**, 5196 (2024) Publisher: Nature Publishing Group.
14. NG Davies, S Flasche, M Jit, KE Atkins, Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nature Ecology & Evolution* **3**, 440–449 (2019) Number: 3 Publisher: Nature Publishing Group.
15. N Mulberry, A Rutherford, C Colijn, Systematic comparison of coexistence in models of drug-sensitive and drug-resistant pathogen strains. *Theoretical Population Biology* **133**, 150–158 (2020).
16. C Colijn, et al., What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of Streptococcus pneumoniae? *Journal of The Royal Society Interface* **7**, 905–919 (2010) Publisher: Royal Society.
17. J Mehtälä, M Antonio, MS Kaltoft, KL O'Brien, K Auranen, Competition Between Streptococcus Pneumoniae Strains: Implications for Vaccine-Induced Replacement in Colonization and Disease. *Epidemiology* **24**, 522–529 (2013) Publisher: Lippincott Williams & Wilkins.
18. E Gjini, C Valente, R Sá-Leão, MGM Gomes, How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *Journal of Theoretical Biology* **388**, 50–60 (2016).
19. LM Childs, et al., Linked within-host and between-host models and data for infectious diseases: a systematic review. *PeerJ* **7**, e7057 (2019) Publisher: PeerJ Inc.
20. SM Eswarappa, S Estrela, SP Brown, Within-Host Dynamics of Multi-Species Infections: Facilitation, Competition and Virulence. *PLOS ONE* **7**, e38730 (2012) Publisher: Public Library of Science.
21. BJZ Quigley, SP Brown, HC Leggett, PD Scanlan, A Buckling, Within-host interference competition can prevent invasion of rare parasites. *Parasitology* **145**, 770–774 (2018).
22. T van der Poll, SM Opal, Pathogenesis, treatment, and prevention of pneumococcal pneumonia. *The Lancet* **374**, 1543–1556 (2009).
23. B Wahl, et al., Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000–15. *The Lancet Global Health* **6**, e744–e757 (2018) Publisher: Elsevier.
24. C Troeger, et al., Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases* **18**, 1191–1210 (2018).
25. WHO, WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance (2024).
26. C Bottomley, A Roca, PC Hill, B Greenwood, V Isham, A mathematical model of serotype replacement in pneumococcal carriage following vaccination. *Journal of The Royal Society Interface* **10**, 20130786 (2013) Publisher: Royal Society.
27. J Corander, et al., Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nature Ecology & Evolution* **1**, 1950–1960 (2017).
28. T Azarian, et al., Frequency-dependent selection can forecast evolution in Streptococcus pneumoniae. *PLOS Biology* **18**, e3000878 (2020) Publisher: Public Library of Science.
29. IM Hastings, Complex dynamics and stability of resistance to antimalarial drugs. *Parasitology* **132** (2006).
30. P Rodrigues, MG M. Gomes, C Rebelo, Drug resistance in tuberculosis—a reinfection model (2006).
31. AH Melnyk, A Wong, R Kassen, The fitness costs of antibiotic resistance mutations. *Evolutionary Applications* **8**, 273–283 (2015).
32. F Luciani, SA Sisson, H Jiang, AR Francis, MM Tanaka, The epidemiological fitness cost of drug resistance in *Mycobacterium tuberculosis*. *Proceedings of the National*

Academy of Sciences **106**, 14711–14715 (2009).
33. MC Maher, et al., The Fitness Cost of Antibiotic Resistance in Streptococcus pneumoniae: Insight from the Field. *PLoS ONE* **7**, e29407 (2012).
34. D Helekal, M Keeling, YH Grad, X Didelot, Estimating the fitness cost and benefit of antimicrobial resistance from pathogen genomic data. *Journal of The Royal Society Interface* **20**, 20230074 (2023) Publisher: Royal Society.
35. GM Knight, et al., The relative fitness of drug-resistant *Mycobacterium tuberculosis* : a modelling study of household transmission in Peru. *Journal of The Royal Society Interface* **15**, 20180025 (2018).
36. P Turner, et al., A Longitudinal Study of Streptococcus pneumoniae Carriage in a Cohort of Infants and Their Mothers on the Thailand-Myanmar Border. *PLOS ONE* **7**, e38271 (2012) Publisher: Public Library of Science.
37. C Turner, et al., High Rates of Pneumonia in Children under Two Years of Age in a South East Asian Refugee Population. *PLOS ONE* **8**, e54026 (2013) Publisher: Public Library of Science.
38. PC Hill, et al., Nasopharyngeal Carriage of Streptococcus pneumoniae in Gambian Infants: A Longitudinal Study. *Clinical Infectious Diseases* **46**, 807–814 (2008).
39. O Abdullahi, et al., Rates of Acquisition and Clearance of Pneumococcal Serotypes in the Nasopharynges of Children in Kilifi District, Kenya. *The Journal of Infectious Diseases* **206**, 1020–1029 (2012).
40. WH Organization, *Pocket Book of Hospital Care for Children: Guidelines for the Management of Common Illnesses with Limited Resources*, Academic Search Complete. (World Health Organization), (2005).
41. S Lehtinen, et al., Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proceedings of the National Academy of Sciences* **114**, 1075–1080 (2017) Publisher: Proceedings of the National Academy of Sciences.
42. M Lipsitch, et al., Estimating Rates of Carriage Acquisition and Clearance and Competitive Ability for Pneumococcal Serotypes in Kenya With a Markov Transition Model. *Epidemiology* **23**, 510–519 (2012).
43. D van der Waaij, JMBd Vries, JECLv der Wees, Colonization resistance of the digestive tract in conventional and antibiotic-treated mice. *The Journal of Hygiene* **69**, 405–411 (1971).
44. H Boureau, T Karjalainen, I Rowland, MHF Wilkinson, L Hartmann, Models to Study Colonisation and Colonisation Resistance. *Microbial Ecology in Health and Disease* **12**, 247–258 (2000) Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08910600050216246.
45. M Lipsitch, et al., Competition among *Streptococcus pneumoniae* for intranasal colonization in a mouse model. *Vaccine* **18**, 2895–2901 (2000).
46. K Auranen, J Mehtälä, A Tanskanen, M S. Kaltoft, Between-Strain Competition in Acquisition and Clearance of Pneumococcal Carriage—Epidemiologic Evidence From a Longitudinal Study of Day-Care Children. *American Journal of Epidemiology* **171**, 169–176 (2010).
47. LJ Hathaway, et al., Capsule Type of Streptococcus pneumoniae Determines Growth Phenotype. *PLOS Pathogens* **8**, e1002574 (2012) Publisher: Public Library of Science.
48. MA Leibold, et al., The metacommunity concept: a framework for multi-scale community ecology. *Ecology letters* **7**, 601–613 (2004).
49. F Bashey, Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140301 (2015) Publisher: Royal Society.
50. T Morel-Journel, et al., Persistence-colonization trade-off and niche differentiation enable the coexistence of E. coli phylogroups (2023) Pages: 2023.08.10.552787 Section: New Results.
51. AR Abruzzo, SD Aggarwal, ME Sharp, GCW Bee, JN Weiser, Serotype-Dependent Effects on the Dynamics of Pneumococcal Colonization and Implications for Transmission. *mBio* **13**, e00158 (2022).
52. V Lázár, O Snitser, D Barkan, R Kishony, Antibiotic combinations reduce Staphylococcus aureus clearance. *Nature* **610**, 540–546 (2022) Publisher: Nature Publishing Group.
53. S Lehtinen, et al., Horizontal gene transfer rate is not the primary determinant of observed antibiotic resistance frequencies in streptococcus pneumoniae. *Science advances* **6**, eaaz6137 (2020).
54. DV McLeod, S Gandon, Understanding the evolution of multiple drug resistance in structured populations. *Elife* **10**, e65645 (2021).
55. M Emons, F Blanquart, S Lehtinen, Temporal Trends in Antibiotic Resistance in Europe, 1998-2019 (2023) Pages: 2023.09.27.23296241.
56. DI Andersson, D Hughes, Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology* **8**, 260–271 (2010).
57. C Davidson-Pilon, lifelines: survival analysis in python. *Journal of Open Source Software* **4**, 1317 (2019).

Supplementary Note 1: The Proportional Hazards assumption

We use survival analysis to analyse longitudinal data on pneumococcal carriage since (a) it can directly model time to event data (time to clearance or time to establishment), and (b) it can easily accommodate censorship in the data. Specifically, we use a survival regression method called the Cox Proportional Hazards model, which is the most generally applicable regression model. The idea of the Cox model is that the log-hazard (risk per unit time) of an event varies linearly with the covariates and the baseline hazard. The Cox model is non-parametric since the baseline hazard function is not fixed unlike in other survival models.

The model relies on the proportional hazards assumption: that the ratio of hazards of a covariate between any two people is a time-independent constant. We test this for our basic Cox model where clearance rate is modelled as dependent on the serotype of the focal strain (Figure S2A). For other models, the time-varying covariates can capture the violations to the rule by accounting for changing covariate values. The test involves confirming that the inferred hazard coefficient (for instance, measured as the scaled Schoenfeld residuals of covariates) do not vary with time (Table S1). We find that serotypes 14, 19A, 19B and 23F do not satisfy the proportional hazards test (at least one of the two p-values < 5%). However, this would not be a limitation for our results since some proportional hazard violations are expected given the large size of our dataset and number of covariates included.

Table S1. Testing the proportional Hazards assumption: Two test statistics are computed for each serotype - (1) KM (Kaplan meier) that checks for violations of the proportional hazards assumption by comparing the survival curves of different groups (stratified by a covariate) to see if they are parallel over time, and (2) Rank based test that evaluates the proportional hazards assumption by looking at the correlation between time and the scaled Schoenfeld residuals of the model

Variable	Test	Test Statistic	p-value	-log2(p)
1	km	0.17	0.68	0.56
	rank	3.40	0.07	3.94
10A	km	0.44	0.51	0.98
	rank	2.60	0.11	3.23
10B	km	0.02	0.89	0.17
	rank	1.34	0.25	2.01
10F	km	0.49	0.48	1.05
	rank	0.67	0.41	1.28
11A	km	0.06	0.80	0.32
	rank	0.38	0.54	0.90
12F	km	0.08	0.78	0.35
	rank	0.15	0.70	0.52
13	km	0.94	0.33	1.59
	rank	0.31	0.58	0.79
14	km	0.38	0.54	0.90
	rank	5.26	0.02	5.52
15A	km	2.21	0.14	2.87
	rank	0.60	0.44	1.18
15B_C	km	1.39	0.24	2.07
	rank	1.56	0.21	2.24
16F	km	2.12	0.14	2.79
	rank	0.88	0.35	1.52
17F	km	0.02	0.89	0.18
	rank	2.34	0.13	2.99
19A	km	4.33	0.04	4.74
	rank	0.01	0.90	0.15
19B	km	6.35	0.01	6.41
	rank	0.83	0.36	1.47
19F	km	2.64	0.10	3.27
	rank	2.35	0.13	3.00
20	km	0.23	0.64	0.66
	rank	0.18	0.67	0.58
21	km	0.30	0.58	0.78
	rank	0.97	0.32	1.63

Variable	Test	Test Statistic	p-value	-log2(p)
22A	km	0.00	0.94	0.08
	rank	1.66	0.20	2.34
22F	km	0.53	0.47	1.10
	rank	0.04	0.85	0.24
23A	km	0.12	0.73	0.45
	rank	0.15	0.70	0.51
23B	km	0.15	0.70	0.52
	rank	0.50	0.48	1.05
23F	km	3.37	0.07	3.91
	rank	5.72	0.02	5.90
24F	km	0.06	0.81	0.31
	rank	1.27	0.26	1.94
28F	km	0.61	0.43	1.21
	rank	0.99	0.32	1.65
29	km	0.08	0.78	0.35
	rank	0.02	0.88	0.19
3	km	0.59	0.44	1.17
	rank	0.25	0.62	0.70
33B	km	0.04	0.84	0.25
	rank	2.56	0.11	3.19
33C	km	0.09	0.76	0.39
	rank	2.68	0.10	3.30
33F	km	0.30	0.89	0.17
	rank	0.25	0.62	0.70
34	km	0.30	0.58	0.78
	rank	0.38	0.54	0.90
35C	km	1.23	0.27	1.90
	rank	0.93	0.33	1.58
35F	km	2.26	0.13	2.91
	rank	0.88	0.35	1.53
38	km	0.28	0.60	0.74
	rank	0.00	0.98	0.02
4	km	0.21	0.65	0.62
	rank	1.06	0.30	1.72
45	km	0.00	0.98	0.03
	rank	1.03	0.31	1.69
46	km	0.01	0.94	0.09
	rank	1.06	0.30	1.72
5	km	0.02	0.90	0.15
	rank	0.09	0.76	0.39
6A_C	km	3.36	0.07	3.90
	rank	0.03	0.87	0.20
6B	km	1.45	0.23	2.13
	rank	0.50	0.48	1.06
7B	km	1.71	0.19	2.39
	rank	0.02	0.90	0.15
7F	km	0.05	0.83	0.27
	rank	1.05	0.30	1.72
8	km	0.06	0.81	0.30
	rank	0.12	0.73	0.46

Variable	Test	Test Statistic	p-value	-log2(p)
9L	km	0.00	0.97	0.04
	rank	0.00	0.96	0.06
9N	km	3.50	0.06	4.02
	rank	0.58	0.44	1.17
9V	km	0.53	0.46	1.11
	rank	0.21	0.65	0.63

Supplementary Note 2: Sensitivity to assumptions about start and end of carriage

Since exact dates of colonisation and clearance of strains cannot be known, we assumed that the colonisation or clearance event happened mid-way between the two sampling points where the potential event occurred. Here, we test the sensitivity of our results on within-host competition to this assumption (Figure S3-A). If the sampling data indicates the start of carriage between T_0 and T_1, we take the time of establishment to be at T_1 - k*(T_1 - T_0), where k=0.25 (Figure S3-A Case A) or 0.75 (Figure S3-A Case B). k=0.5 corresponds to results from the main text. Similarly if clearance occurs between times T_2 and T_3, then the time of establishment is T_2 + k*(T_3 - T_2). We analyse the results on within-host competition for these values of k.

We find that the presence of within-host competitors always increases the rate of clearance of strains for k values of 0.25, 0.5 and 0.75 (Figure S3-Bi). The hazard coefficients of presence-of-competitor are 0.28 (CI=0.22 to 0.36) for k=0.25, 0.28 (CI=0.21 to 0.35) for k=0.5 and 0.25 (CI=0.18 to 0.32) for k=0.75. The effect of serotype on within-host competition is tested using a model with interaction coefficients between serotype and presence-of-competitor, and regressing the serotype coefficients with the interaction coefficients (Figure 3B). The slopes of this regression are always negative for k=0.25 to 0.75 (Figure S3-Bii). The slope is -0.41 (CI=-0.60 to -0.21) for k=0.25, -0.40 (CI=-0.60 to -0.20) for k=0.5 and -0.42 (CI=-0.60 to -0.23) for k=0.75. Similarly, the effect of presence of competitors on rate of establishment is also not sensitive to the values of k tested (Figure S3-Biv). The hazard coefficient of presence-of-competitor on establishment are -0.80 (CI=-0.90 to -0.69) for k=0.25, -0.79 (CI=-0.90 to -0.68) for k=0.5 and -0.75 ((CI=-0.86 to -0.64) for k=0.75.

The priority effect that we observe on the clearance of co-colonized strains is sensitive to the values of k tested (Figure S3-Biii). The results of this analysis are significant for k=0.5 and 0.75. The difference in clearance rates between a resident and non-resident competitor is not significantly different from each other when k=0.25, but both have a positive hazard of clearance. The absence of an effect for k=0.25 might be driven by scenarios where there is consecutive loss and gain of different serotypes in a host (i.e., serotype replacement). When two different serotypes are detected in two consecutive samples in the data, then k=0.25 (and not 0.5 and 0.75) corresponds to the assumption that the earlier serotype is lost before establishment of the new serotype such that no co-colonisation occurs. In the case of k=0.5 and 0.75, the new serotype would be assumed to replace the pre-existing serotype shortly after co-colonisation. So the effect of a resident serotype may be acting primarily on invading strains that result in serotype replacement.

Supplementary Note 3: Effect of antibiotics on establishment rates

In the main text, we analysed the effect of antibiotic consumption on clearance rates of resistant and susceptible strains. Here, we report the effects of antibiotic consumption on establishment rates. Since comparison of establishment rates of resistant versus susceptible strains is confounded by the total prevalence of resistance in the population, we can only infer the effects of antibiotics on susceptible strains, captured by the variable *effective-drug-use*.

Similar to the corresponding clearance model, we perform survival analysis on the times to establishment of serotypes using *presence-of-competitor*, *effective-drug-use*, and *ineffective-drug-use* as covariates, and age of infant and serotype as stratifying variables. The hazard coefficient of *effective-drug-use* is 0.00 (CI=-0.09 to 0.09) (Figure S5A). We find no signal for an effect of drug consumption on the rate of establishment of drug-susceptible strains, even when analysed for amoxicillin / ampicillin separately (Figure S5C). However, the coefficient value of *effective-drug-use* is sensitive to the assumption of length of drug-use efficacy of 7 days, and *effective-drug-use* is associated with a positive value when this length is assumed to be 15 days (more in Note 6).

Supplementary Note 4: Effect of different antibiotics on clearance rates

In this section, we quantify how antibiotic usage (captured by the effect sizes of *effective-drug-use* and *ineffective-drug-use*) affects clearance rates across all drug treatments in the dataset.

We use a Cox proportional hazards model with serotype, *presence-of-competitor*, *effective-drug-use*, and *ineffective-drug-use* included as covariates (similar to the model in main text). The covariates *effective-drug-use*, and *ineffective-drug-use* are defined with respect to a single antibiotic at a time. The results of this analysis are shown in Figure S5B. Only four antibiotics are shown here; the model does not converge well for the others since there are only few recorded instances of use of those drugs. For amoxicillin/ampicillin, the hazard coefficient of *effective-drug-use* is 0.55 (CI=0.41 to 0.68) and of *ineffective-drug-use* is 0.34

(CI=0.13 to 0.54). This corresponds to an increase in clearance rate by 73% (CI=51% to 97%) of susceptible strains and by 40% (CI=14% to 73%) of resistant strains. Similar to the analysis across all drugs, we see that susceptible strains are cleared faster by amoxicillin/ampicillin than resistant strains. For all other antibiotics, 95% confidence intervals of their effect sizes overlap zero.

Supplementary Note 5: Antibiotic-blind analyses of the costs of resistance

In the main text, we quantified the fitness costs of resistance, i.e., the effect of resistance on the clearance rate of strains under no antibiotic exposure conditions (using the covariate *resistance-without-drug*). Here, we look at whether resistance has any overall fitness effects *without* stratifying the data by drug consumption. In other words, we quantify how *resistance* affects fitness of a strain while being blind to whether the strain experiences antibiotic pressure. We test this for three fitness components - clearance and establishment rates, and within-host competition between strains.

First, we test whether resistant strains have a higher clearance rate in our dataset than susceptible strains. We run a time-varying proportional hazards model that includes serotype of focal strain, *presence-of-competitor* and resistance status of focal strain as covariates. The resistance variable is used in two ways (in separate analyses): (1) Resistance to antibiotic X, for all X, and (2) Resistance to any antibiotic. We do not include ceftriaxone resistance, since there are only two episodes of colonisation in the data that are ceftriaxone resistant. We find that the hazard coefficient of resistance is variable from positive to negative depending on the antibiotic tested (Figure S6A). For 3 antibiotics, the hazard is negative and significant, and for 2 antibiotics, is negative with CI overlapping zero. For these five antibiotics, the results suggest that resistance is associated with a longer duration of carriage. We find a signal indicating a clearance cost in resistance to chloramphenicol, which has a significant positive hazard of 0.24 (CI=0.02 to 0.47). This corresponds to a 27.4% higher clearance rate of chloramphenicol-resistant strains compared to chloramphenicol-susceptible strains.

We then look at the interaction term between *presence-of-competitor* and resistance in the above model, for all seven antibiotics. The interaction coefficients measure the hazard of clearance due to resistance (compared to susceptibility) in the presence of within-host competitors, thereby capturing the cost of resistance on within-host competition. None of the interaction terms between resistance and *presence-of-competitor* are significantly different from zero (Figure S6B). In this analysis, we find no signal for a cost of resistance on within-host competitive ability during clearance.

In the above analyses, we looked at how resistance of the *focal strain* affects its clearance. Next, we analyse how resistance status of a within-host competitor affects clearance of the focal strain. To do this, we define the time-dependent variables *presence-of-resist-competitor* and *presence-of-sens-competitor*, which indicate whether the within-host competitor, if present, is resistant or susceptible respectively. We use a time-varying proportional hazards model with the following covariates: serotype of focal strain, *presence-of-resist-competitor*, and *presence-of-sens-competitor*. The difference in the values of coefficients between *presence-of-resist-competitor* and *presence-of-sens-competitor*, signifies the effect of resistance on a competitor's ability to induce clearance of the focal strain. As before, we repeat the analysis for all seven antibiotics independently, and by assuming resistance to *any* antibiotic. The hazard coefficients of *presence-of-resist-competitor* and *presence-of-sens-competitor* are not significantly different from each other for any of the antibiotics tested (Figure S6C). We do not find any signal for an effect of within-host competitors' resistance on the clearance rate of the focal strain.

Finally, we analyse the effect of resistance on establishment rate of strains. We test whether the competitive effect faced by an establishing strain depends on whether the resident strain is resistant or susceptible to antibiotics. This is achieved using survival analysis on 'time to establishment' data using the following covariates: *presence-of-resist-competitor* and *presence-of-sens-competitor*, with age of host and serotype of incoming strain used to stratify the baseline hazard rate. The difference in hazard coefficients between *presence-of-resist-competitor* and *presence-of-sens-competitor* is found to be non-significant (Figure S7D) for all resistances tested, thus excluding evidence for a cost of resistance on establishment rates.

Supplementary Note 6: Assumptions about drug consumption in hosts

We assumed in our analyses that drug consumption, where included, has an effect in hosts for 7 days. We test the sensitivity of our results to this assumption by varying the assumed duration of drug efficacy to 5, 10 and 15 days (Figure SS9).

Antibiotic consumption is always associated with an increased rate of clearance of resistant and sensitive strains for all drug-use lengths tested (positive coefficients of *effective-* and *ineffective-drug-use* in Figure 5A and Figure S9A). The difference between the two hazard coefficients however is sensitive to this assumption: the hazard of *effective-drug-use* is significantly higher than that of *ineffective-drug-use* when the drug-use duration is 10 days or higher. At drug-use length of 10 days, the hazard coefficients are 0.62 (CI=0.52 to 0.72) for *effective-drug*-use and 0.19 (CI=0.02 to 0.38) for *ineffective-drug-use*. The difference between these two coefficients signify the fitness benefit of resistance on rate of clearance of strains. Resistance is associated with a significant fitness benefit during antibiotic use when we assume the length of drug-use to be 10 or more days.

The result on the effect of drug use on rate of establishment is sensitive to the lengths of drug-use tested (Figure S5A and Figure S9B). The strength of the coefficient of *effective-drug-use* is positive and significant for the case of 15 days, which corresponds to a higher establishment rate of sensitive strains. This may be because, by assuming 15 days, the model captures more cases of post-antibiotic-use establishment of sensitive strains. If so, then this coefficient captures the post-treatment repopulation of the host niche rather than the effect of treatment itself.

Figure	Model	AIC	AIC - min(AIC)
Figure 3A	Model (in main text)	61620.21	0
	Model without serotype	63042.06	1421.85
Figure 3B	Model	89897.86	0
Figure 3C	Model	61491.22	0
	Model without serotype	63006.52	1515.30
Figure 3D	Model	13315.5	0
Figure 5A	Model	62180.7	0
	Model without serotype	62976.80	796.10
	Model without competitor	62852.89	672.2
	Model - both	63040.70	860.0
Figure 5B	Model	62162.50	0
	Model without serotype	62937.9	775.40
Figure 5C	Model	13329.30	0

Table S2. Partial AIC values for Cox models from analyses in the main text.

All results corresponding to measuring the costs of resistance are non-sensitive to the assumptions about length of drug-use. These include - the effect of resistance on clearance rate (Figure S9 C), cost of resistance on competitive ability during clearance (Figure S9 C) and establishment (Figure S9 D), and the costs of resistance measured for individual resistances (Figure S9 E-F).

Supplementary Note 7: Model selection and parsimony estimations

We use Akaike Information criterion (AIC) to calculate model parsimony. AIC is a measure of relative quality of statistical models that balances the risk of over-fitting by adding too many variables versus the risk of under-fitting by using an oversimplified model.

For a model with k independent variables and a maximum likelihood estimate of L, the value of AIC is given by the equation:

$$AIC = 2(k - ln(L)) \quad \textbf{(S1)}$$

For each model, we extracted the AIC of the model, and compared to that of all models with a subset of covariates that can be used to analyze the effects we study in the model. In Cox models, we calculate the partial AIC since the model maximizes a partial likelihood function that focuses on the relative hazard rates between individuals without requiring knowledge of the baseline hazard function. A lower partial AIC corresponds to a more parsimonious model. The values of partial AIC tested for different models are listed in Table S2.

Supplementary-Figures

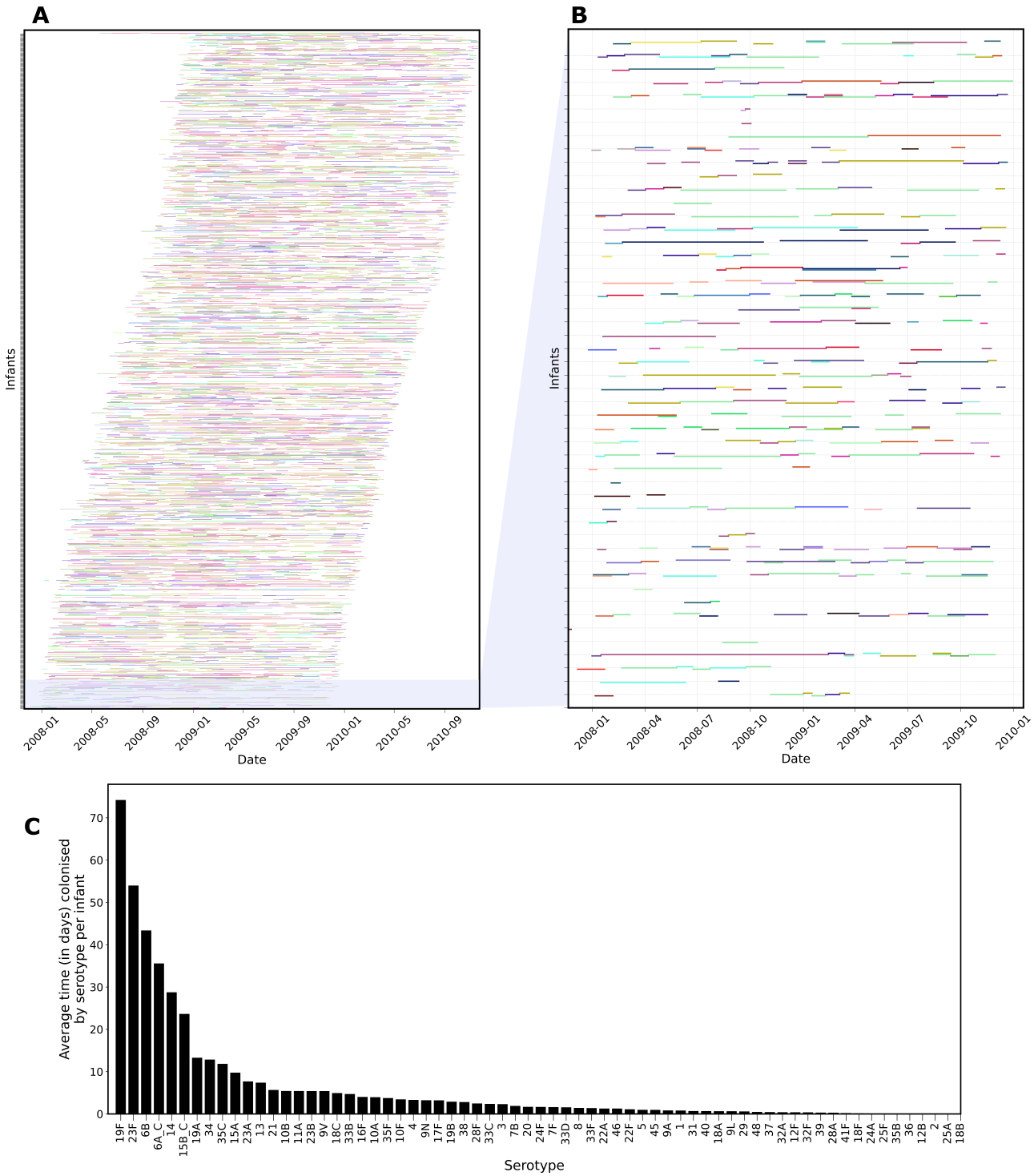

Figure S1. Brief overview of the data used to run the survival analyses. Each row along the vertical axis corresponds to an individual infant. Each coloured horizontal line represents the carriage of a strain in the host along time (shown on the X axis). Each colour represents a unique serotype. The serotypes in each row are slightly offset vertically to aid the visualization of overlapping carriages within an infant. (A) The durations of carriage of all observed strains in all infants used for the analyses. (B) The first 50 infants (from the bottom of A) are zoomed into to show carriage of strains in individual infants. (C) Serotype distribution in the dataset. Each bar corresponds to the sum of all observed durations of carriage of the serotype divided by the total number of infants.

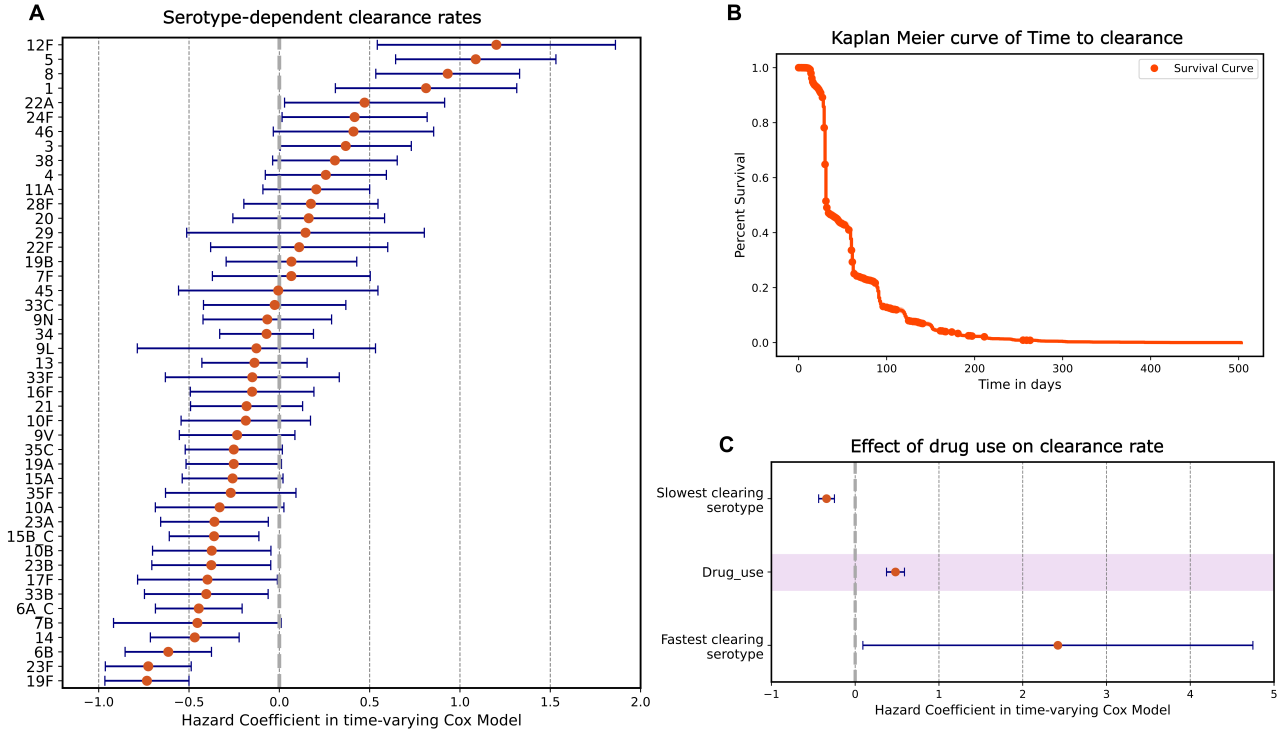

Figure S2. Basic survival models to describe pneumococcal dynamics (A) Clearance rate (and thus the duration of carriage) varies with the serotype of the strain. (B) For the model in A, the survival curve that summarizes the duration of carriage of all strains in our data. On the Y axis is plotted the fraction of episodes of carriage that have a duration of carriage that lasts at least as long as the value corresponding to the X axis. (C) Does antibiotic consumption affect the clearance rate of strains? The Drug use covariate has a positive coefficient of 0.46 (CI=0.36 to 0.57) indicating increased clearance rates associated with antibiotic use.

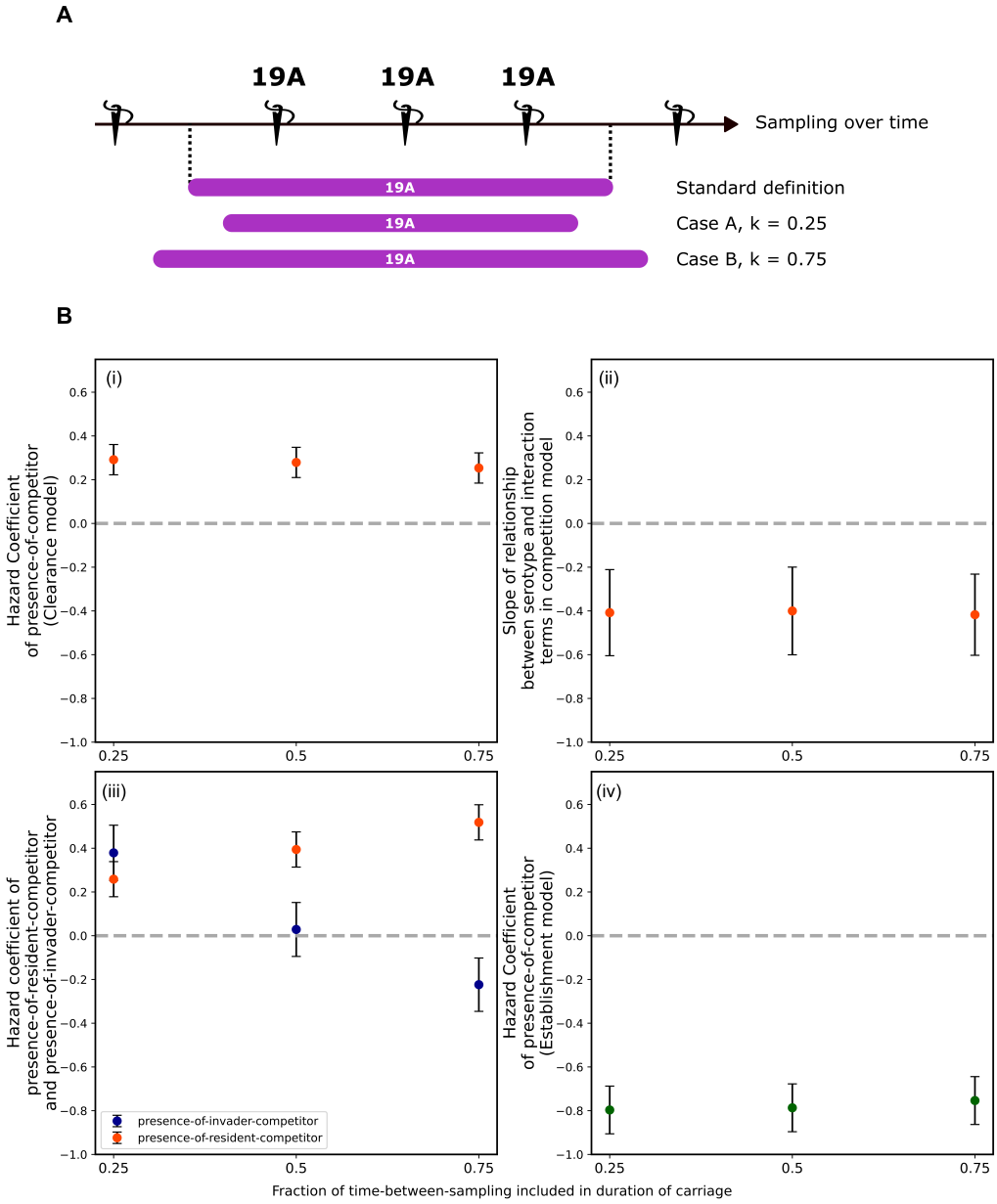

Figure S3. Testing sensitivity to assumptions about start and end of carriage. (A) An example schematic is shown where a strain with serotype 19A is detected in 3 consecutive samples. The standard definition of duration of carriage assumes start and end of carriage to be half-way between two sampling dates. We modify this assumption in Case A and B where these start and end dates to be 25% or 75% of the duration between two sampling dates respectively. (B) We test the effects of assumptions in A on the results on within-host competition. The effects of within-host competition on clearance and establishment are not sensitive to the assumption about start and end dates of carriage, except in one case. The resident-priority effect in clearance (as in Figure 1C) between co-colonized strains is not significant in the case of 25% (Bottom left panel).

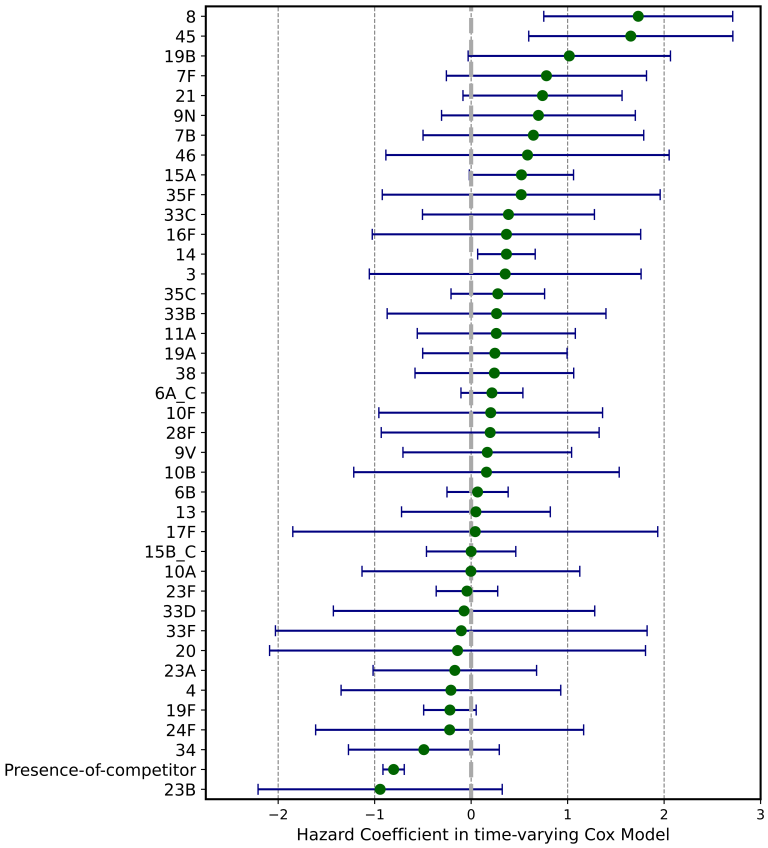

Figure S4. Does serotype of the resident strain affect rate of establishment of the incoming strains? The model includes resident serotypes as covariates to look at their effects on establishment. Serotypes can have a range of different competitive abilities, ranging from preventing new establishment to facilitating new establishment.

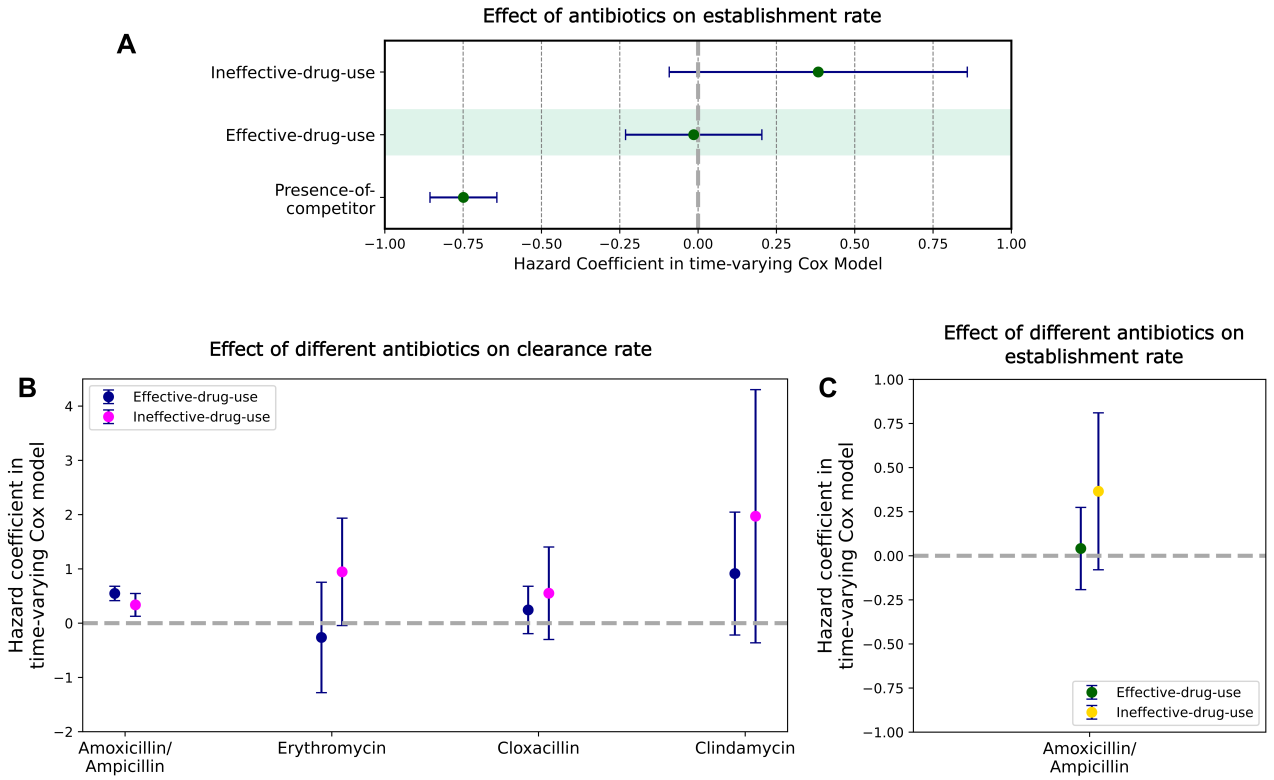

Figure S5. (A) Effect of drug consumption on establishment rates. The coefficient of *effective-drug-use* captures the effect of antibiotics on the establishment rate of drug-susceptible strains. (B,C) Effects of drug consumption on clearance rates (B), and establishment rates (C), analysed for each antibiotic in separate models. The hazard coefficients of *effective-drug-use* and *ineffective-drug-use* are shown for different antibiotics. For other drugs not shown in the figure, there is not enough data for convergence of the Cox model.

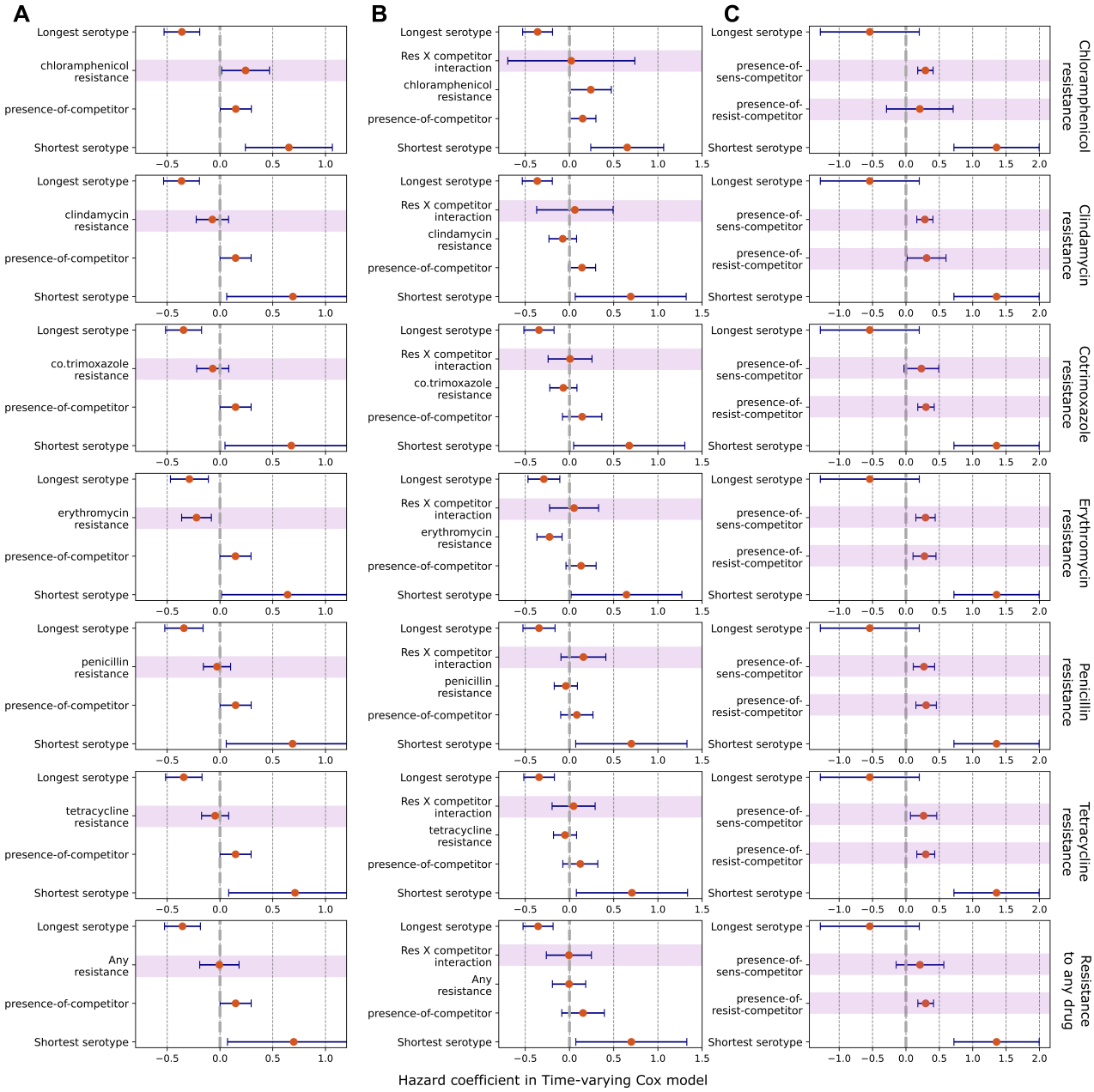

Figure S6. Effects of resistance on rate of clearance of serotypes (A) Does resistance to antibiotics affect the clearance rate of focal strain? The resistance variable tracks this for different antibiotics. (B) Does resistance to antibiotic affect clearance in the presence of a competitor? Here, the interaction term between resistance and presence-of-competitor tracks this effect for different drugs. (C) Does the resistance status of the co-colonizer strain affect clearance rate of the focal strain? We infer this based on the difference between coefficients of presence-of-sens-competitor and presence-of-resist-competitor for different drugs.

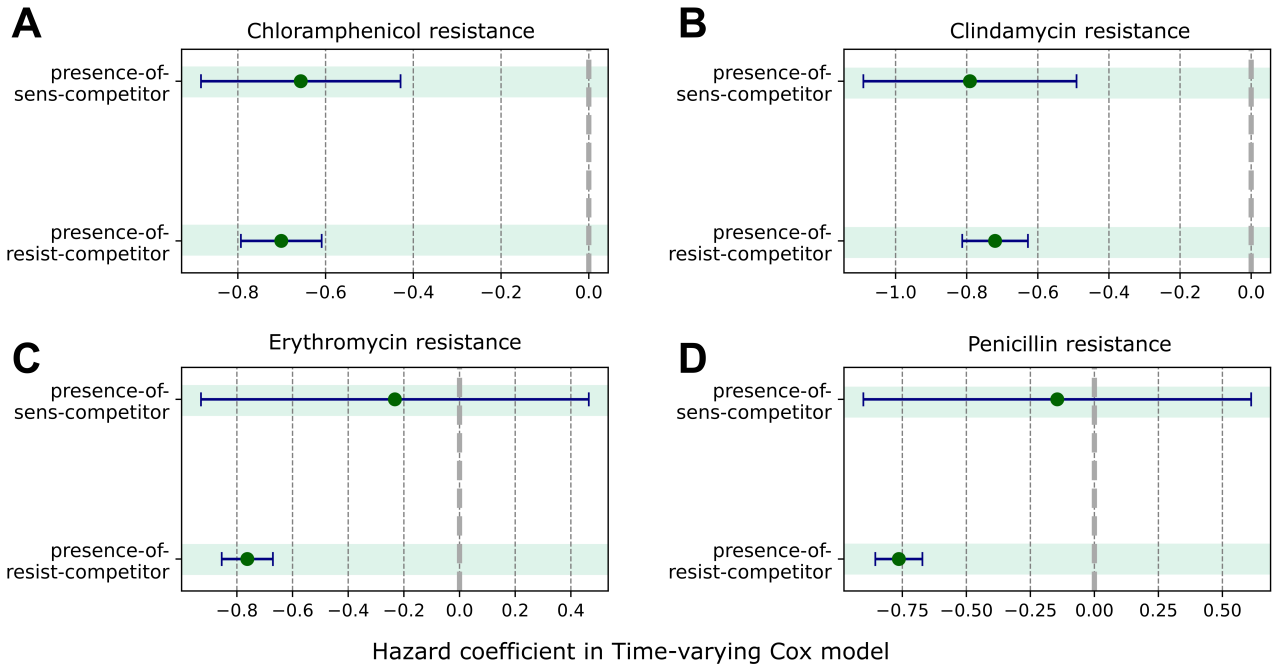

Figure S7. Effects of resistance of within-host competitor on the establishment rate of focal strain, tested for resistance to four drugs in A-D. For the other resistances, the data was insufficient for model convergence. The difference in coefficients between presence-of-sens-competitor and presence-of-resist-competitor are shown for (A) chloramphenicol resistance, (B) Clindamycin resistance, (C) Erythromycin resistance, and (D) Penicillin resistance.

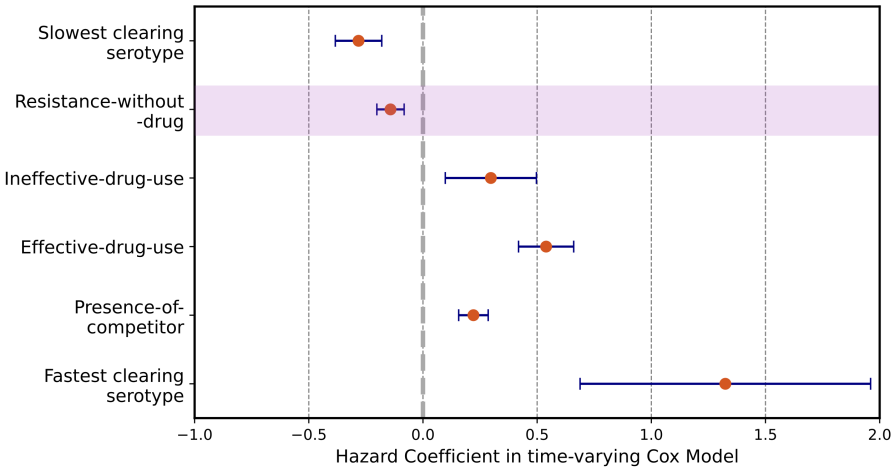

Figure S8. Does antibiotic resistance affect clearance rate of strains? The covariate resistance-without-drug captures the effect of being resistant on rate of clearance. The negative coefficient indicates that resistant strains are associated with a lower clearance rate compared to susceptible strains.

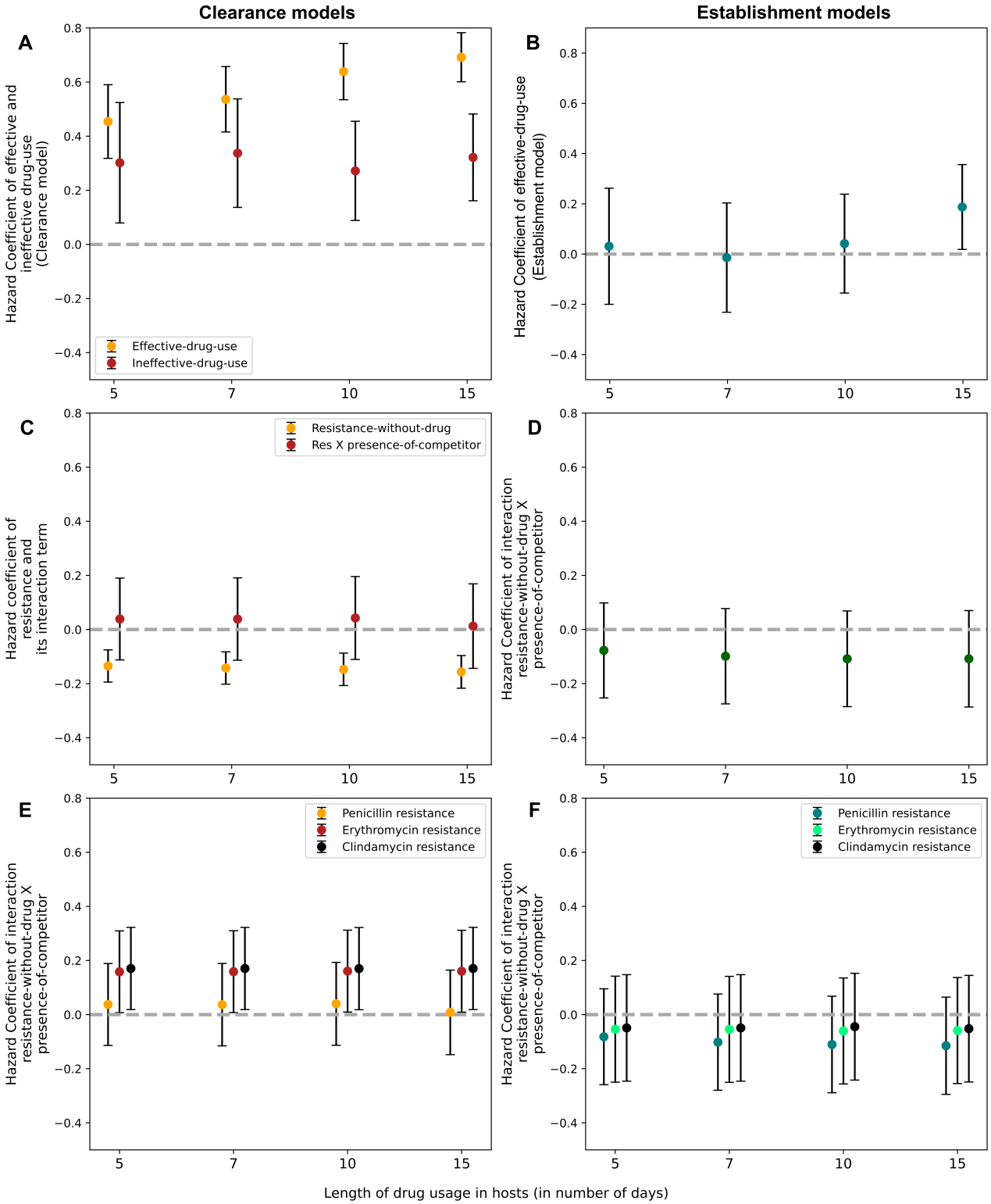

Figure S9. Testing sensitivity of results to drug-use duration. A,C,E are based on models of time to clearance, and B,D,F are based on models of time to establishment. (A) The hazards coefficients of effective-drug-use and ineffective-drug-use on clearance rate are always positive. At 10 days and 15 days, the difference in hazards between effective-drug-use and ineffective-drug-use is significant. (B) The hazards coefficients of effective-drug-use and ineffective-drug-use on establishment rate are always non-significant, except at 15 days, when effective-drug-use has a positive coefficient. (C) The effect of resistance on clearance in the absence of drugs, and the effect of resistance on within-host competitive ability, are not sensitive to the durations of drug-use tested. (D) The within-cost of resistance during establishment is always negative and non-significant. (E,F) Within-host cost of resistance to different antibiotics during clearance (E) and establishment (F) are non-sensitive to duration of drug use.