1 Title and authors

Title: Multiple mechanisms of strain competition are needed to explain pneumococcal serotype-specific antibiotic resistance patterns

1.0.1 Author list

Hannah C. Lepper (1), Emma Pujol-Hodge (2), Jada Hackman (1,3), Kevin van Zandvoort (4), Kate Mellor-Wright (3), Oliver Lorenz (3), Stephanie Lo (3,5), Duc Anh Dang (6), Hung Thai Do (7), Gerry Tonkin-Hill (8, 9), Stephen D. Bentley (3), Lay Myint Yoshida (10), Stefan Flasche (4, 11), Nicholas G. Davies (4), Katherine E. Atkins (1, 4)

1.0.2 Affiliations

1. Usher Institute, University of Edinburgh, UK
2. Public Health Scotland, Edinburgh, UK
3. Parasites and Microbes, Wellcome Sanger Institute, UK
4. 4. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, UK
5. Department of Life Sciences, University of Bath, UK
6. National Institute of Hygiene and Epidemiology, Vietnam
7. Pasteur Institute of Nha Trang, Vietnam
8. University of Melbourne, Australia

9. Peter MacCallum Cancer Centre, Australia

10. Nagasaki University, Japan

11. Charité Centre for Global Health, Germany

2 Abstract

Background: Disease caused by antibiotic resistant pneumococcal strains is an important public health concern. The impact of public health interventions depend on the complex ecology governing serotype and antibiotic resistant strain dynamics, for which we currently lack mechanistic explanations. Ee evaluated whether multi-serotype transmission models of antibiotic resistant pneumococcal strains can recapitulate empirical resistance data, thus providing a plausible explanation for how antibiotic-sensitive and resistant strains coexist at intermediate frequencies, both between and within serotypes.

Methods: We extracted a global dataset of pneumococcal carriage prevalence in healthy, unvaccinated children by serotype and penicillin susceptibility from a previously-published systematic review. Next, we developed a suite of multi-serotype individual-based models that have previously been shown to plausibly explain coexistence in single serotype models, incorporating transmission and clearance rate differences between serotypes and serotype-specific immunity. Finally, we calibrated the suite of models to the relative serotype prevalence and overall resistance frequency drawn from the extracted dataset.

Results: Overall resistance frequency varied considerably between studies, but we found a positive association between frequency of a serotype and resistance frequency within that serotype, and a high degree of coexistence of sensitive and resistant strains within serotypes. Each individual-based model predicted different serotype-specific resistance frequencies, but combining models was necessary to capture all features of the

empirical data.

Conclusions: Serotype-specific resistance patterns may give important clues to understand fundamental antibiotic resistance epidemiology. We show that combining multiple independently plausible mechanisms can capture resistant pneumococcal carriage, but additional data are needed to determine the strength of these individual mechanisms within these combinations.

3 Introduction

Antibiotic-resistant bacterial infections are estimated to kill around 1.14 million people annually (Murray et al., 2022; Naghavi et al., 2024). Designing interventions that effectively reduce drug resistant infections is therefore a global priority (Vekemans et al., 2021). Predicting the impact of an intervention aiming to reduce resistance requires first understanding which factors drive the frequency with which bacteria are resistant to antibiotics in the absence of an intervention. However, finding such a mechanistic explanation has been a fundamental challenge (Lipsitch et al., 2009; Davies et al., 2021). Moreover, for many bacterial species that cause substantial antibiotic resistant deaths, such as *Streptococcus pneumoniae* and *Escherichia coli*, explaining the stable intermediate frequency of resistance becomes harder, because the species have multiple serotypes which themselves differ in their level of antibiotic resistance (Andrejko et al., 2022; Weerdenburg et al., 2023). As we currently lack a mechanistic understanding of how antibiotic resistance is stably maintained across bacterial serotypes, we have limited ability to predict the impact of interventions designed to reduce disease and deaths from antibiotic resistance.

In *S. pneumoniae*, which accounts for an estimated 10% of deaths due to antibiotic resistant infections (Murray et al., 2022; Naghavi et al., 2024), resistance is more frequent in the most prevalent serotypes, which cause the most invasive disease (Lo et al., 2019; Andrejko et al., 2022; Song et al., 2012). There are over 100 different serotypes (Ganaie

et al., 2020), each of which has a different polysaccharide capsule that elicits serotype-specific immune responses (Hausdorff et al., 2005). Serotype also differ in traits such as carriage duration which could impact selection for resistance (Lehtinen et al., 2017; Lees et al., 2017).

Some serotypes are more likely to cause invasive disease than others, and pneumococcal conjugate vaccines (PCVs) protect against transmission and disease from a subset of virulent serotypes (Gallagher et al., 2025; Lewnard and Hanage, 2019). PCV has been proposed as a key intervention for reducing antibiotic resistance, and overall resistance frequency in pneumococcal carriage and infection has been observed to reduce after vaccine introduction (Andrejko et al., 2021). However, resistance frequency to some antibiotics has also increased in non-PCV serotypes (Lo et al., 2019; Andrejko et al., 2021), suggesting complex dynamics may be at play. To capture pneumococcal resistance epidemiology and predict the impact of interventions, it is important to be able to explain how resistance is selected for at the serotype level.

Several microbial ecological mechanisms could explain the co-occurrence of sensitive and resistant strains of the same bacterial species in host populations (Lehtinen et al., 2017; Davies et al., 2019; Blanquart et al., 2018; Krieger et al., 2020; Davies et al., 2021; Colijn et al., 2010). However, we still do not have a coherent understanding of how these mechanisms interact with serotype dynamics to generate empirical patterns of pneumococcal resistance. To address this, we built a dataset of global serotype-specific resistance patterns in pneumococcal carriage in healthy children before PCV introduction, and developed a suite of multi-serotype individual-based mathematical models for each mechanism. By comparing predictions of each model individually and in combination to our dataset, we assessed the role of each mechanism in shaping serotype-level selection for resistance.

4 Results

4.1 Serotype-specific resistance patterns in pneumococcal carriage studies

To build a dataset of global serotype-specific resistance patterns, we systematically identified studies reporting serotype-specific prevalence and penicillin susceptibility among studies included in an existing systematic review Andrejko et al. (2021). Of the 558 studies screened, we found 12 studies that met all inclusion criteria (S. Fig 3.), covering 9,977 children from 12 countries between 1995 and 2019 (Supplementary Materials). The average pneumococcal carriage prevalence across all samples was 35.5%, with study prevalences ranging from 14.1% in Singapore in 2008 to 78.4% in Norway in 2006, which the authors identify as an unusually high prevalence for this setting Vestrheim et al. (2008). In total, 50 serotypes were identified, with 41.4% of all isolates resistant or non-susceptible to penicillin, with individual study frequencies ranging from 3.7% in Sweden in 2005 to 93.5% in Jordan in 2019.

We applied mixed-effects binomial regression models to describe serotype-specific resistance patterns in the data. Under the best-fitting model, we found a significant (p=0.03) positive association between serotype frequency rank and resistance frequency (7% increase in the odds of non-susceptibility per unit increase in rank) (S.Table 4). However, over half of the variance in the observed data (52%) was explained by differences between studies, indicating that other unobserved factors associated with region affect resistance frequency of a serotype. Serotype identity itself only explained 4% of variation in the data while serotype rank as a fixed effect explained 2% (S.Table 4).

Coexistence of resistant and sensitive strains of the same serotype was observed in all studies. In the study with the most samples, Espinosa-de los Monteros et al. (2007), 94.1% of serotypes with more than five or more isolates had an intermediate resistance frequency (between 5 and 95%). However, in many studies there were low number of

isolates or a low resistance frequency, making serotype-specific resistance frequency within that study uncertain (Fig 1).

4.2 Serotype-specific resistance patterns predicted by mathematical models

To evaluate our ability to recapitulate these empirical data, we developed a suite of four individual-based transmission models for each of the coexistence mechanisms that can plausibly reproduce non-serotype specific resistance frequency identified in Davies et al. (2021) (Table 1). We then used data on relative serotype prevalence and overall resistance frequency from the extracted dataset to calibrate the models.

To capture a single epidemiological setting, we selected the largest dataset extracted from the systematic review, Espinosa-de los Monteros et al. (2007), which reported 2777 samples from 12 states in Mexico, extracting the frequencies of the 10 most prevalent serotypes and the overall penicillin resistance rate (65.7%) (2B).

To test the performance of these calibrated models, we then compared each model output with three key qualitative observations generated from the entire systematic review dataset: 1) overall population-level coexistence of sensitive and resistant strains (of any serotype); 2) a positive association between serotype frequency and resistance frequency; 3) and high within-serotype coexistence, i.e. a high proportion of serotypes maintaining an intermediate resistance frequency.

Each individual model generated a range of stable intermediate frequencies of resistance with increasing treatment rate when incorporating multiple serotypes, consistent with results from the respective single-serotype models (S. Fig 4). However, each model generated a different serotype-specific resistance pattern, and no single model could simultaneously capture both the positive serotype frequency-resistance frequency association and intermediate proportions of resistance frequency within serotypes. Under the

treatment diversity model, every serotype had the same resistance frequency, with increasing variance with decreasing serotype rank (Fig 3A, Table 2). Consistent with previous studies (Lehtinen et al., 2017), the pathogen diversity model predicted that serotypes above the carriage duration threshold are all resistant, and those below are all sensitive (Fig 3A, Table 2). Both the treatment and growth competition models predicted that serotypes tend to have close to either 0% or 100% resistance frequency, with intermediate resistance frequency within a serotype being rare (Fig 3A, Table 2). Whereas resistance frequency was positively associated with serotype prevalence in the treatment competition model, the growth competition model resulted in the opposite pattern (Fig 3A, Table 2). We investigated why there was a lack of intermediate resistance frequencies within serotypes in the within-host models with a simplified, 2-serotype version of the model and invasion experiments, which indicated that immune competition can prevent both strains of the same serotype from coexisting (described in Supplementary Materials).

We then created two-, three- and four-way combinations of each mechanism to test whether observed serotype-specific patterns were obtainable when multiple drivers of resistant-sensitive strain competition were operating simultaneously (Fig 4). Adding the treatment diversity mechanism to any other model increased within-serotype coexistence, without affecting the predicted relationship between serotype and resistance frequency (Table 2, Fig 4A). The treatment competition or the pathogen diversity models could still capture positive serotype frequency-resistance frequency relationships when combined either together or with the treatment diversity model (Table 2, Fig 4A). However the growth competition model performed poorly in all 2-way combinations, predicting a negative relationship between serotype and resistance frequencies (Fig 4A). Combining the treatment diversity, treatment competition, and pathogen diversity mechanisms also agreed with observed patterns (Table2, Fig 4B). We did not combine the treatment and growth competition models in a 2-way combination because this would have resulted in two different resistance costs in the same model, and would have required additional fitting to

be meaningful.

In the treatment diversity model, sub-groups of the population with different treatment rates mix assortatively. To test of whether this differential mixing between subpopulations alone was enough to generate intermediate resistance frequencies in the combined models, we combined the treatment competition model and the treatment diversity model without any variation in treatment rate between groups (S. Fig 5). The serotype-specific resistant patterns under this combination were comparable to treatment competition model acting alone.

Finally, we tested whether a small within-host growth advantage for sensitive strains in addition to all other mechanisms was compatible with observed serotype-specific resistance patterns using a 4-way combination of all models. Within this model, there are two costs to resistance: a transmission cost and a within-host growth cost. We added 1/10 and 1/2 of the growth advantage sensitive strains needed to maintain the target resistance frequency when the growth competition was acting alone. Under this combination of models, both the empirically-observed positive association between serotype frequency and resistance frequency and within-serotype coexistence was maintained (Table 2, Fig 4C, S. Fig 6).

5 Discussion

Using empirical carriage data and individual-based multi-serotype models, we identified sets of mechanisms of sensitive-resistant strain competition that can capture pneumococcal resistance epidemiology. Of the mechanisms studied here, we show that both host population transmission structure and treatment heterogeneity were needed to explain coexistence of sensitive and resistant strains of the same serotypes; that serotype-specific carriage duration or within-host competition (or both) could explain differences in resistance frequency by serotype; and that differential within-host growth of sensitive

and resistant strains may not be strong enough to play a role in population-level transmission dynamics. We highlight what further empirical data and analyses that are needed to further hone in on the strength of each mechanism that would best explain resistance dynamics.

Although within-host competition models have been shown to capture empirical resistance frequencies in single-serotype models, we show that serotype-level coexistence is lost or depleted when these models are extended to include multiple serotypes, so that these co-colonisation models alone could not replicate carriage data. We suggest this loss of serotype-level coexistence is due to host adaptive immunity, which blocks same-serotype co-colonisation that is required to generate the negative-frequency dependent selection which maintains the intermediate resistance frequency in the absence of multiple serotypes. Our argument is supported by our previous modelling work, which found more extensive intermediate resistance frequencies within serotypes when immunity is gained on clearance of a serotype rather than adaptive immunity being gained following colonisation (Davies et al., 2019). However, recent immunological evidence suggests that serotype-specific adaptive immunity is gained on acquisition, not clearance (Ramos-Sevillano et al., 2019). Our results highlight the potential importance of interactions between immune-mediated negative-frequency dependent selection (NFDS) and other NFDS processes that act on loci that impact strain fitness on pathogen diversity. In this case, two nested diversifying selection processes resulted in depletion rather than maintenance of within-serotype diversity.

Serotype-level co-existence seen in empirical data can be regained in the within-host competition co-colonisation model when it is combined with host population structure (via the treatment diversity model). The treatment diversity model permits coexistence of sensitive and resistant strains by creating niches for each serotype in the host population and thus overriding the effects of adaptive immunity. On its own, a population structure model was insufficient to explain why serotypes would have different resistance frequencies.

However, the combined model captured both the association between serotype carriage prevalence and resistance frequency and the within-serotype co-existence observed empirically. Integrating population structure into the pathogen diversity model worked in a similar way: by itself each mechanism explained one of the components of the empirical data, while combined they explained both. It is well established in ecological theory that niche structure (for example spatial structure) can promote coexistence of one trait (Chesson, 2000), but we are not aware of applications to explain nested diversity (such as sensitive and resistant strains coexisting within serotypes) in pathogens.

Our results show that faster transmission of some serotypes could explain the positive association between serotype prevalence and resistance frequency. This finding is consistent with previous studies (Davies et al., 2019). That serotypes or pneumococcal genomes with longer carriage duration have a higher resistance frequency is well established in theory and carriage studies (Lehtinen et al., 2017; Andrejko et al., 2022; Lees et al., 2017). However, to our knowledge, there has not been an analysis of loci associated with faster transmission and resistance frequency in pneumococcal genomes. PCV changes the transmission rates of serotypes, and it has been shown theoretically that this could impact resistance dynamics post-PCV (Davies et al., 2021). Empirically it is observed that resistance frequency increases in non-vaccine serotypes (Lo et al., 2019; Andrejko et al., 2021), which supports the theory that transmission could affect strength of selection for resistance. It is therefore crucial to show how influential transmission rates are on serotype-level resistance selection to accurately predict resistant carriage changes after vaccine introduction.

Our results suggest that a within-host growth resistance cost alone cannot capture patterns in pneumococcal resistance data, nor when combined only with carriage duration diversity or host population structure by treatment group. The growth competition model on its own and in 2-way combinations predicts a negative relationship between serotype prevalence and resistance frequency, which has not been empirically observed. However,

sensitive strains can outcompete resistant strains due to growth rate-related fitness costs in vitro (Melnyk et al., 2015). How are these two observations compatible? Here we propose one solution by including a weak effect of differential growth alongside other strain competition mechanisms and a transmission resistance cost. When the growth differential is small, treatment usually occurs before within-host competition has resulted in loss of resistant strains, so that this process has little impact on the overall force of infection of sensitive and resistant strains, and is therefore not driving NFDS for sensitive strains. An alternative consideration that we have not addressed is that serotypes themselves differ in their within-host growth rate, as it has been shown that serotypes that are more competitive in vitro tend to be more resistant in carriage data (Andrejko et al., 2022). To what extent mathematical models of infectious disease need to account for both within-host and between-host scales to explain population-level trends is an ongoing question (Mideo et al., 2008; Childs et al., 2019).

Although we show that combining models is essential to capture all qualitative patterns in serotype-specific resistance data, we do not know how strongly each mechanism could be driving resistance patterns. It is likely that several different combinations of model strengths could explain observed data equally well, and that the parameters controlling the strength of each driver are unidentifiable. Larger datasets on community carriage that report serotype and antibiotic susceptibility would reduce uncertainty and permit more validation. In addition, data on co-colonisation would narrow down the set of likely combinations by adding data specific to the within-host model mechanisms, while controlled observations after an intervention may also provide insights (Davies et al., 2021).

6 Methods

6.1 Search strategy, study selection criteria, and data analysis

To identify studies reporting serotype-specific prevalence and penicillin susceptibility, we restricted our search to those included in a recent systematic review and meta-regression of carriage and invasive disease (Andrejko et al., 2021) that had already undergone comprehensive screening and quality assessment. We further restricted studies to those reporting serotype-specific prevalence and penicillin susceptibility in nasopharyngeal carriage isolates collected from children, prior to widespread introduction of pneumococcal conjugate vaccines (PCVs) in the population. We then excluded studies where participants had a mean age of less than 10 years old, and where children had been recruited because they were immunocompromised or had had a diagnosed respiratory condition.

For each study, we extracted the total number of samples taken, the total number of pneumococcal positive samples, the fraction of susceptible (i.e. sensitive) and non-susceptible isolates, the minimum inhibitory concentration (MIC) threshold used to indicate non-susceptibility and/or resistance in the study, country, region, information on age of cohort, and percentage of children vaccinated (as some children were vaccinated in the private market in the absence of national immunisation programmes).

Within each study, we then classified each pneumococcal positive isolate according to serotype or serogroup where serotype was not available, and penicillin susceptibility (with non-susceptible isolates broken down into intermediate or resistant isolates, where possible). From these data, for each serotype, we calculated 1) the serotype frequency (the fraction of isolates in the study belonging to that serotype), 2) the frequency rank (in decreasing order), and 3) the non-susceptibility frequencies (the fraction of isolates below the resistance cut off).

To understand the general association between non-susceptibility to penicillin and other variables, we fitted four binomial generalised linear mixed-effects models to the

full dataset: 1) the null model, (intercept-only), 2) the serotype rank model (serotype frequency rank as a fixed effect), 3) the serotype class model (serotype identity as a random effect), and 4) the serotype class and rank model (serotype frequency rank as a fixed effect and serotype identify as a random effect). All models included study random effects to account for variation in susceptibility across settings, and an observation-level random effect, to deal with overdispersion. We used Akaike Information Criteria (AIC) to determine the best-fitting model.

6.2 Mathematical models

6.2.1 Description

We developed a suite of four individual-based transmission models for each of the coexistence mechanisms previously shown to plausibly explain resistance frequency identified in Davies et al. (2021), and based on the individual-based pneumococcal transmission model described in Davies et al. (2019). A full description of the models is provided in the Supplementary Materials. In brief, we first constructed a base model for L serotypes, with one sensitive and one resistant strain for each subtype ($L * 2 = M$ strains in total). Each serotype is denoted $l \in \{1, \ldots, L\}$ and each strain is denoted $j \in \{1, \ldots, M\}$. An individual, $i \in \{1, \ldots, N\}$, has a continuous colonisation state with respect to the proportion of the colonisation niche taken up by each strain (f_{ij}), a total colonisation status which is always 0 or 1 ($F_i = \sum f_{ij}$), and a serotype-specific immune state (m_l) of 0 or 1. We characterised the following events: 1) transmission of strain j of serotype l and simultaneous gain of immunity to serotype l, which occurs in immunologically naive hosts at rate $\kappa^{F_i}\lambda_j/N$, where κ is the relative rate of secondary colonisation to primary colonisation, and λ_j is the strain-specific force of infection, based on a serotype-specific transmission rate β_l and the sum of all colonisations of strain j in the population; 2) natural clearance, in which all strains are simultaneously lost, which occurs at rate μ, which is not serotype-

specific in the base model; 3) immune waning, in which all immunity is lost, which occurs at rate ϕ; and 4) death or ageing of the individual, in which the colonisation status of individual i is reset, representing death or ageing out of the tracked population br and replacement with a fully susceptible newborn. As demonstrated in Davies et al. (2019), this base model is structurally neutral with respect to sensitive and resistant strains of each serotype (Lipsitch et al., 2009), so that all coexistence between sensitive and resistant is explained by the inputted mechanisms. Events are modelled as inhomogeneous Poisson processes. We only track colonisations in children (average age of 10 years old) to parsimoniously capture the majority of transmission events, which occur between children (Qian et al., 2022; Van Zandvoort et al., 2024), and to match carriage data.

Next, we incorporated four coexistence mechanisms into separate versions of this base model: treatment diversity (TD), pathogen diversity (PD), treatment competition (TC), and growth competition (GC) (Table 1, Supplementary Materials). We add antibiotic treatment events, which remove sensitive strains from hosts at rate τ for all models, although we do not expect clearing pneumococcal colonisation to be the target of antibiotic treatments. We also included a cost of resistance, as a reduction in transmission of the resistant strain for TD, PD, and GC models (c), relative to the susceptible strain, and as a benefit to growth of sensitive strain in co-colonisations in the GC model (ω). Co-colonisation is not possible in TD and PD models ($\kappa = 0$).

We also combined the models. When the treatment or growth competition models were combined with the pathogen diversity model, hosts may be colonised with more than one serotype and therefore subject to multiple serotype-specific clearance rates. Therefore, we adapted the model so that the clearance rate was a weighted average of the serotype-specific clearance rates of the serotypes with which they are currently colonised, $\sum f_i \mu_l$. We obtained 20 repeated simulations of each model individually, and for 2-, 3-, and 4-way combinations.

6.2.2 Model parameterisation and calibration

First, we calibrated the transmission rates of each serotype in the base model without sensitive and resistant strain dynamics to the relative serotype-specific prevalence in the target dataset. We used a model emulator approach (Perumal and Zyl, 2025) that fitted a Gaussian Process model to the serotype prevalence ratios from model simulations obtained for different transmission rate ratios. We then used the emulator model to estimate the transmission rate ratios needed to obtain the serotype prevalence ratios observed in the data (see Supplementary Materials for full details).

Next, we fit each individual model to an overall resistance frequency by estimating the cost of resistance in each model. We used Bayesian Optimization (Pelikan, 2005) to find the resistance cost that minimised the distance between observed and model obtained resistance frequency (see Supplementary Materials for full details). For the PD model, it is possible to exactly estimate the resistance cost thresholds at which each serotype will be resistant, based on the treatment rate and the carriage duration of each serotype. We generated a resistance frequency for the mid-point between each threshold through model simulation, and selected the resistance cost that gave the most similar resistance frequency to the data.

Parameters that were not related to the strain competition mechanism (including average carriage duration and duration of immunity) were drawn from the literature and previous model fits (S.Table 2) and did not differ between models.

6.3 Software, data, and code availability

We conducted data and statistical analyses using R v4.3.2 (R Core Team, 2023) and the lme4 (Bates et al., 2015), DHARMa (Hartig, 2024) and sjPlot (LÃijdecke, 2024) packages. The individual-based models were implemented in Julia v1.11 (Bezanson et al., 2017), with the calibration and fitting using the GaussianProcesses (Fairbrother et al., 2022) and

BayesianOptimization (Brea, 2021) packages.

7 Acknowledgements

KEA, HCL, JH and EP-H were supported by the Wellcome Trust [219797/Z/19/Z]. KvZ was supported by the Bill and Melinda Gates Foundation [OPP1139859]. SF was supported by the Einstein Foundation Berlin as an Einstein BUA Strategic Professor [EPP-BUA-2022-697].

Kynan Delaney for helpful statistical discussions.

Mechanism	Description	Mechanism of coexistence	Sources
Treatment diversity (TD)	The host population is split into groups with different treatment rates, with more transmission within than between groups.	Resistant strains are selected for more strongly in some groups than others, and is slowly imported into low treatment groups from high treatment groups	Krieger et al., 2018, Blanquart et al., 2018
Pathogen diversity (PD)	Each serotype has a different clearance rate, so that carriage duration is longer for some serotypes than others	Serotypes with longer duration are more exposed to treatment and have stronger selection for resistance, so that some serotypes are resistant and others are sensitive	Lehtinen et al., 2017
Treatment competition (TC)	Sensitive and resistant strains can co-colonise hosts. Treatment clears sensitive strains only, allowing the resistant strain to outcompete co-colonising sensitive strains.	Resistant strain is selected for more strongly when it is rare as it is more likely to be a co-coloniser with the more common sensitive strains, generating negative frequency dependent selection	Davies et al., 2019
Growth competition (GC)	Strains can co-colonise, but sensitive strains can outcompete resistant strains within the host before treatment because they have a faster growth rate.	Sensitive strain is selected for more strongly when it is rare as it is more likely to be a co-coloniser with the more common resistant strains, generating negative frequency dependent selection	Davies et al., 2019

Table 1: **Description of each modelled mechanism.** Drawn from (Davies et al., 2019).

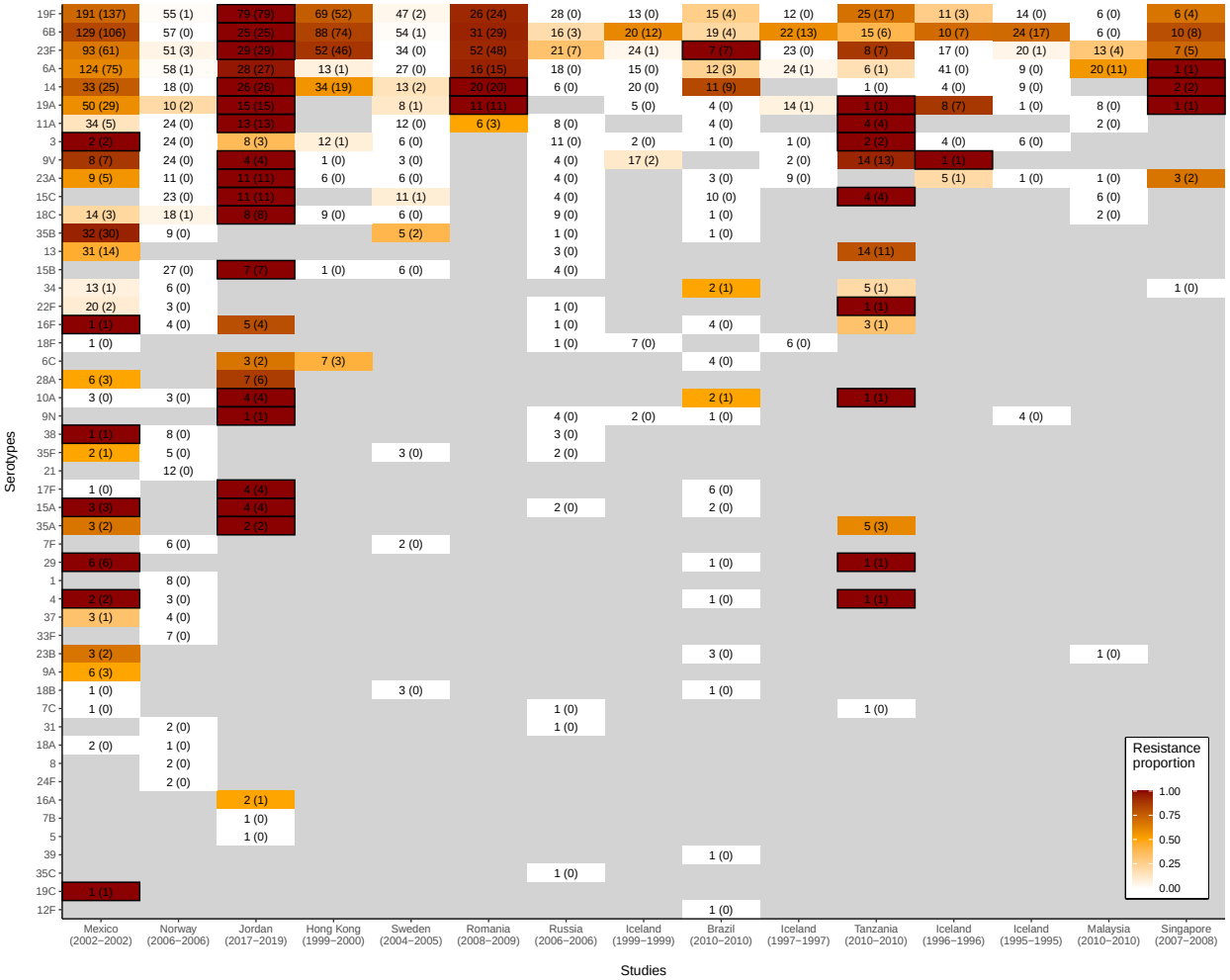

Figure 1: **Resistance frequency varied by study, serotype, and serotype rank in the empirical data.** Total *S. pneumoniae* isolates and non-susceptible isolates (in brackets) for each study and serotype. The colour of the tiles indicates the relative non-susceptible frequency for the serotype, with black-outlined boxes indicating cases of 100% non-susceptible frequency. Studies on the x-axis are in decreasing order of total number of isolates analysed, while serotypes on the y-axis are in decreasing order of total prevalence.

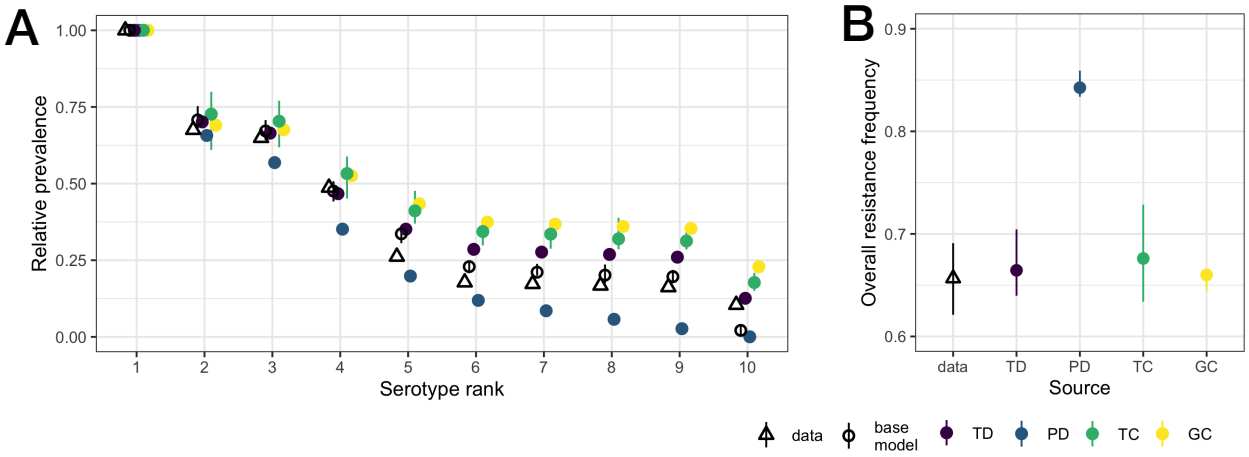

Figure 2: **model fitting results**

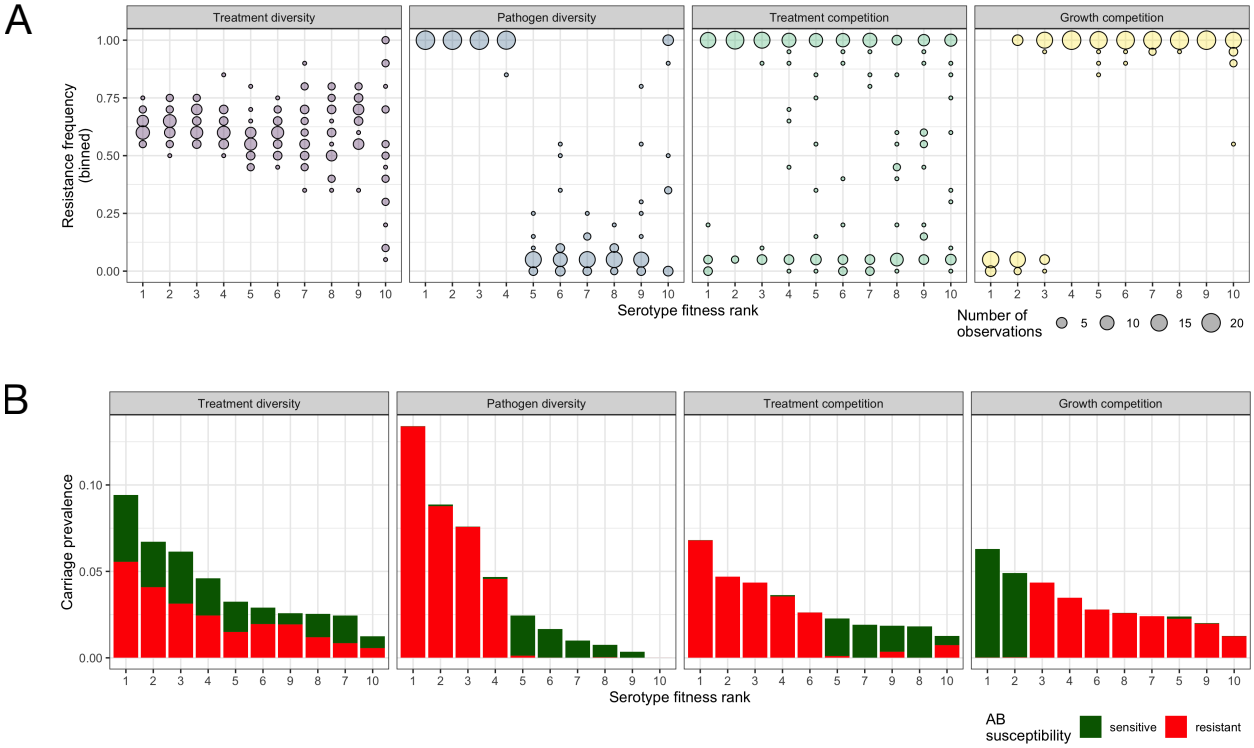

Figure 3: **Each mechanism predicts a different serotype-specific resistance pattern.** A) Variation in resistance frequency of each hypothetical serotype from 20 model simulations. Resistance frequency (y axis) is binned, with the size of the circle indicating the number of repeated model simulations that fell within that bin. B) Example carriage prevalence of the 10 most prevalent serotypes for a single simulation of each model. Red indicates the proportion of samples of a given serotype that were non-susceptible to penicillin and green represents the proportion that were sensitive.

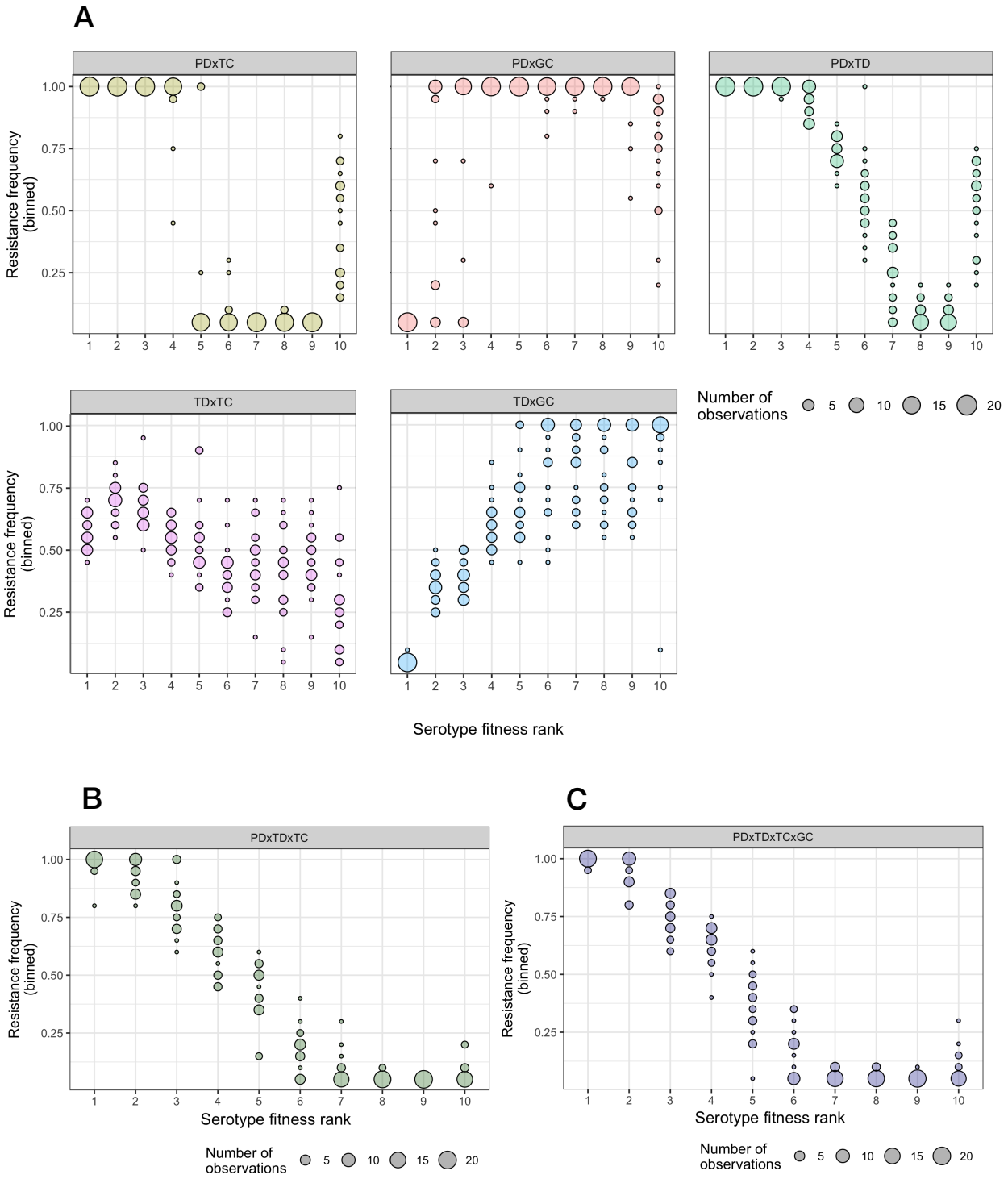

Figure 4: **A range of serotype-specific resistance patterns is possible when combining mechanisms** Variation in resistance frequency of each hypothetical serotype from 20 model simulations of combinations of the four models. Resistance frequency (y axis) is binned, with the size of the circle indicating the number of repeated model simulations that fell within that bin. A) two-way combinations between the models, B) three-way, and C) four-way combinations of the models.

Model	Captures positive relationship between serotype prevalence and resistance frequency?	Captures high levels of within-serotype coexistence?
Individual models:		
Treatment diversity (TD)	No (0.03, -0.10 to 0.17)	Yes (0.98, 0.96 to 1.00)
Pathogen diversity (PD)	Yes (0.62, 0.52 to 0.70)	No (0.20, 0.16 to 0.25)
Treatment competition (TC)	Yes (0.22, 0.09 to 0.36)	No (0.26, 0.20 to 0.31)
Growth competition (GC)	No (-0.38, -0.58 to -0.25)	No (0.02, 0.00 to 0.03)
Combined models:		
PDxTD	Yes (0.86, 0.82 to 0.89)	Yes (0.52, 0.47 to 0.56)
PDxTC	Yes (0.66, 0.58 to 0.73)	No (0.12, 0.12 to 0.24)
PDxGC	No (-0.04, -0.18 to 0.10)	No (0.20, 0.15 to 0.24)
TDxTC	Yes (0.64, 0.56 to 0.72)	Yes (0.99, 0.96 to 1.00)
TDxGC	No (-0.80, -0.85 to -0.74)	Yes (0.69, 0.65 to 0.74)
PDxTDxTC	Yes (0.89, 0.86 to 0.92)	Yes (0.50, 0.46 to 0.54)
PDxTDxTCxGC	Yes (0.87, 0.83 to 0.90)	Yes (0.49, 0.43 to 0.54)

Table 2: **Only combined models can capture all qualitative characteristics in the empirical data** Comparison of serotype-specific resistance pattern characteristics in the data and the model. Spearman's Rho gives the correlation between serotype fitness rank and resistance frequency rank within each repeated model simulation, with 95% confidence intervals. Within-serotype coexistence proportion gives the average proportion of serotypes that had intermediate resistance frequencies (between 5 and 95%) across model simulations, with 95% quantiles.

References

K. Andrejko, B. Ratnasiri, W. P. Hausdorff, R. Laxminarayan, and J. A. Lewnard. Antimicrobial resistance in paediatric Streptococcus pneumoniae isolates amid global implementation of pneumococcal conjugate vaccines: a systematic review and meta-regression analysis. *The Lancet Microbe*, 2(9):e450–e460, Sept. 2021. ISSN 26665247. doi: 10.1016/S2666-5247(21)00064-1. URL https://linkinghub.elsevier.com/retrieve/pii/S2666524721000641.

K. Andrejko, B. Ratnasiri, and J. A. Lewnard. Association of Pneumococcal Serotype With Susceptibility to Antimicrobial Drugs: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases*, 75(1):131–140, Aug. 2022. ISSN 1058-4838, 1537-6591. doi: 10.1093/cid/ciab852. URL https://academic.oup.com/cid/article/75/1/131/6380293.

D. Bates, M. MÃďchler, B. Bolker, and S. Walker. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.

J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017. doi: 10.1137/141000671. URL https://epubs.siam.org/doi/10.1137/141000671. Publisher: SIAM.

F. Blanquart, S. Lehtinen, M. Lipsitch, and C. Fraser. The evolution of antibiotic resistance in a structured host population. *Journal of The Royal Society Interface*, 15(143): 20180040, June 2018. ISSN 1742-5689, 1742-5662. doi: 10.1098/rsif.2018.0040. URL https://royalsocietypublishing.org/doi/10.1098/rsif.2018.0040.

J. Brea. BayesianOptimization. jl. *Julia package version 0.2*, 5, 2021.

P. Chesson. Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology and Systematics*, 31(1):343–366, Nov. 2000.

ISSN 0066-4162. doi: 10.1146/annurev.ecolsys.31.1.343. URL https://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.343.

L. M. Childs, F. El Moustaid, Z. Gajewski, S. Kadelka, R. Nikin-Beers, J. W. Smith, Jr, M. Walker, and L. R. Johnson. Linked within-host and between-host models and data for infectious diseases: a systematic review. *PeerJ*, 7:e7057, June 2019. ISSN 2167-8359. doi: 10.7717/peerj.7057. URL https://peerj.com/articles/7057.

C. Colijn, T. Cohen, C. Fraser, W. Hanage, E. Goldstein, N. Givon-Lavi, R. Dagan, and M. Lipsitch. What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae* ? *Journal of The Royal Society Interface*, 7(47):905–919, June 2010. ISSN 1742-5689, 1742-5662. doi: 10.1098/rsif.2009.0400. URL https://royalsocietypublishing.org/doi/10.1098/rsif.2009.0400.

N. G. Davies, S. Flasche, M. Jit, and K. E. Atkins. Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nature Ecology & Evolution*, 3(3): 440–449, Feb. 2019. ISSN 2397-334X. doi: 10.1038/s41559-018-0786-x. URL https://www.nature.com/articles/s41559-018-0786-x.

N. G. Davies, S. Flasche, M. Jit, and K. E. Atkins. Modeling the effect of vaccination on selection for antibiotic resistance in *Streptococcus pneumonia e*. *Science Translational Medicine*, 13(606):eaaz8690, Aug. 2021. ISSN 1946-6234, 1946-6242. doi: 10.1126/scitranslmed.aaz8690. URL https://www.science.org/doi/10.1126/scitranslmed.aaz8690.

L. E. Espinosa-de los Monteros, V. Jiménez-Rojas, F. Aguilar-Ituarte, M. Cashat-Cruz, A. Reyes-López, R. Rodríguez-Suárez, P. Kuri-Morales, R. Tapia-Conyer, and D. Gómez-Barreto. Aislamientos de s. pneumoniae en niños sanos de estancias infantiles en 12 estados de méxico. *Salud Pública de México*, 49(4):249–255, 2007.

J. Fairbrother, C. Nemeth, M. Rischard, J. Brea, and T. Pinder. Gaussianprocesses. jl: A nonparametric bayes package for the julia language. *Journal of Statistical Software*, 102:1–36, 2022.

K. E. Gallagher, F. Odiwour, C. Bottomley, J. Ojal, A. Adamu, E. Muthumbi, E. W. Kagucia, L. L. Hammitt, S. Massora, B. SigaÃžque, A. ChaÃžque, L. Vilanculos, J. R. Verani, M. Da Gloria Carvalho, A. Von Gottberg, J. Kleynhans, S. A. Madhi, C. P. Olwagen, G. Mackenzie, R. Salaudeen, R. Gierke, M. Kobayashi, S. Pelton, I. Yildirim, S. Thomas, A. Tunali, M. Farley, T. D. Swarthout, A. KalizangâĂŹoma, R. S. Heyderman, N. French, Y. Choi, N. Andrews, S. Ladhani, E. Miller, and J. A. G. Scott. Serotype-specific pneumococcal invasiveness: a global meta-analysis of paired estimates of disease incidence and carriage prevalence, Mar. 2025. URL http://medrxiv.org/lookup/doi/10.1101/2025.03.03.25323223.

F. Ganaie, J. S. Saad, L. McGee, A. J. Van Tonder, S. D. Bentley, S. W. Lo, R. A. Gladstone, P. Turner, J. D. Keenan, R. F. Breiman, and M. H. Nahm. A New Pneumococcal Capsule Type, 10D, is the 100th Serotype and Has a Large *cps* Fragment from an Oral Streptococcus. *mBio*, 11(3):e00937–20, June 2020. ISSN 2161-2129, 2150-7511. doi: 10.1128/mBio.00937-20. URL https://journals.asm.org/doi/10.1128/mBio.00937-20.

F. Hartig. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models, 2024. URL https://cran.r-project.org/package=DHARMa.

W. P. Hausdorff, D. R. Feikin, and K. P. Klugman. Epidemiological differences among pneumococcal serotypes. *Lancet Infectious Diseases*, 5(2):83–93, 2005.

M. S. Krieger, C. E. Denison, T. L. Anderson, M. A. Nowak, and A. L. Hill. Population structure across scales facilitates coexistence and spatial heterogeneity of antibiotic-resistant infections. *PLOS Computational Biology*, 16(7):e1008010,

July 2020. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1008010. URL https://dx.plos.org/10.1371/journal.pcbi.1008010.

J. A. Lees, N. J. Croucher, D. Goldblatt, J. Parkhill, C. Turner, P. Turner, and S. D. Bentley. Genome-wide identification of lineage and locus specific variation associated with pneumococcal carriage duration. *eLife*, 2017. doi: https://doi.org/10.7554/eLife.26255.

S. Lehtinen, F. Blanquart, N. J. Croucher, P. Turner, M. Lipsitch, and C. Fraser. Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proceedings of the National Academy of Sciences*, 114(5):1075–1080, Jan. 2017. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1617849114. URL https://pnas.org/doi/full/10.1073/pnas.1617849114.

J. A. Lewnard and W. P. Hanage. Making sense of differences in pneumococcal serotype replacement. *The Lancet Infectious Diseases*, 19(6):e213–e220, June 2019. ISSN 14733099. doi: 10.1016/S1473-3099(18)30660-1. URL https://linkinghub.elsevier.com/retrieve/pii/S1473309918306601.

M. Lipsitch, C. Colijn, T. Cohen, W. P. Hanage, and C. Fraser. No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics*, 1(1): 2–13, Mar. 2009. ISSN 17554365. doi: 10.1016/j.epidem.2008.07.001. URL https://linkinghub.elsevier.com/retrieve/pii/S1755436508000029.

S. W. Lo, R. A. Gladstone, A. J. van Tonder, J. A. Lees, M. du Plessis, R. Benisty, N. Givon-Lavi, P. A. Hawkins, J. E. Cornick, B. Kwambana-Adams, P. Y. Law, P. L. Ho, M. Antonio, D. B. Everett, R. Dagan, A. von Gottberg, K. P. Klugman, L. McGee, R. F. Breiman, S. D. Bentley, A. W. Brooks, A. Corso, A. Davydov, A. Maguire, A. Pollard, A. Kiran, A. Skoczynska, B. Moiane, B. Beall, B. Sigauque, D. Aanensen, D. Lehmann, D. Faccone, E. Foster-Nyarko, E. Bojang, E. Egorova, E. Voropaeva, E. Sampane-Donkor, E. Sadowy, G. Bigogo, H. Mucavele, H. BelabbÃĺs, I. Diawara, J. MoÃŕsi, J. Verani,

J. Keenan, J. N. Nair Thulasee Bhai, K. M. Ndlangisa, K. Zerouali, K. L. Ravikumar, L. Titov, L. De Gouveia, M. Alaerts, M. Ip, M. C. de Cunto Brandileone, M. Hasanuzzaman, M. Paragi, M. Nurse-Lucas, M. Ali, N. Elmdaghri, N. Croucher, N. Wolter, N. Porat, Ã. K. Eser, P. E. Akpaka, P. Turner, P. Gagetti, P.-E. Tientcheu, P. E. Carter, R. Mostowy, R. Kandasamy, R. Ford, R. Henderson, R. Malaker, S. Shakoor, S. C. Grassi Almeida, S. K. Saha, S. Doiphode, S. A. Madhi, S. Devi Sekaran, S. Srifuengfung, S. Obaro, S. C. Clarke, S. A. Nzenze, T. Kastrin, T. J. Ochoa, V. Balaji, W. Hryniewicz, and Y. Urban. Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-PCV13 era: an international whole-genome sequencing study. *The Lancet Infectious Diseases*, 19(7): 759–769, July 2019. ISSN 14733099. doi: 10.1016/S1473-3099(19)30297-X. URL https://linkinghub.elsevier.com/retrieve/pii/S147330991930297X.

D. LÃijdecke. sjPlot: Data Visualisation for Statistics in Social Science, 2024. URL https://cran.r-project.org/package=sjPlot.

A. H. Melnyk, A. Wong, and R. Kassen. The fitness costs of antibiotic resistance mutations. *Evolutionary Applications*, 8(3):273–283, Mar. 2015. ISSN 1752-4571, 1752-4571. doi: 10.1111/eva.12196. URL https://onlinelibrary.wiley.com/doi/10.1111/eva.12196.

N. Mideo, S. Alizon, and T. Day. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends in Ecology & Evolution*, 23 (9):511–517, Sept. 2008. ISSN 01695347. doi: 10.1016/j.tree.2008.05.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0169534708002188.

C. J. L. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. Kashef Hamadani, E. A. P. Kumaran, B. McManigal, S. Achalapong, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews,

A. Aravkin, E. Ashley, F.-X. Babin, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, J. A. Berkley, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. CastaÃśeda-Orjuela, V. Chansamouth, S. Chaurasia, S. ChiurchiÃź, F. Chowdhury, R. Clotaire Donatien, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. J. Day, M. De Luca, K. Dokova, A. Dramowski, S. J. Dunachie, T. Duong Bich, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, C. Garcia, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, Y. Hsia, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, A. W. J. Jenney, M. Khorana, S. Khusuwan, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, K. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, A. Manoharan, F. Marks, J. May, M. Mayxay, N. Mturi, T. Munera-Huertas, P. Musicha, L. A. Musila, M. M. Mussi-Pinhata, R. N. Naidu, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, T. J. Ochoa, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, P. Ounchanum, G. D. Pak, J. L. Paredes, A. Y. Peleg, C. Perrone, T. Phe, K. Phommasone, N. Plakkal, A. Ponce-de Leon, M. Raad, T. Ramdin, S. Rattanavong, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, V. D. Rosenthal, K. E. Rudd, N. Russell, H. S. Sader, W. Saengchan, J. Schnall, J. A. G. Scott, S. Seekaew, M. Sharland, M. Shivamallappa, J. Sifuentes-Osornio, A. J. Simpson, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Tigoi, C. Turner, P. Turner, H. R. Van Doorn, S. Velaphi, A. Vongpradith, M. Vongsouvath, H. Vu, T. Walsh, J. L. Walson, S. Waner, T. Wangrangsimakul, P. Wannapinij, T. Wozniak, T. E. M. W. Young Sharma, K. C. Yu, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek, and M. Naghavi. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, 399(10325): 629–655, Feb. 2022. ISSN 01406736. doi: 10.1016/S0140-6736(21)02724-0. URL https://linkinghub.elsevier.com/retrieve/pii/S0140673621027240.

M. Naghavi, S. E. Vollset, K. S. Ikuta, L. R. Swetschinski, A. P. Gray, E. E. Wool, G. Robles Aguilar, T. Mestrovic, G. Smith, C. Han, R. L. Hsu, J. Chalek, D. T. Araki, E. Chung, C. Raggi, A. Gershberg Hayoon, N. Davis Weaver, P. A. Lindstedt, A. E. Smith, U. Altay, N. V. Bhattacharjee, K. Giannakis, F. Fell, B. McManigal, N. Ekapirat, J. A. Mendes, T. Runghien, O. Srimokla, A. Abdelkader, S. Abd-Elsalam, R. G. Aboagye, H. Abolhassani, H. Abualruz, U. Abubakar, H. J. Abukhadijah, S. Aburuz, A. Abu-Zaid, S. Achalapong, I. Y. Addo, V. Adekanmbi, T. E. Adeyeoluwa, Q. E. S. Adnani, L. A. Adzigbli, M. S. Afzal, S. Afzal, A. Agodi, A. J. Ahlstrom, A. Ahmad, S. Ahmad, T. Ahmad, A. Ahmadi, A. Ahmed, H. Ahmed, I. Ahmed, M. Ahmed, S. Ahmed, S. A. Ahmed, M. A. Akkaif, S. Al Awaidy, Y. Al Thaher, S. O. Alalalmeh, M. T. AlBataineh, W. A. Aldhaleei, A. A. S. Al-Gheethi, N. B. Alhaji, A. Ali, L. Ali, S. S. Ali, W. Ali, K. Allel, S. Al-Marwani, A. Alrawashdeh, A. Altaf, A. B. Al-Tammemi, J. A. Al-Tawfiq, K. H. Alzoubi, W. A. Al-Zyoud, B. Amos, J. H. Amuasi, R. Ancuceanu, J. R. Andrews, A. Anil, I. A. Anuoluwa, S. Anvari, A. E. Anyasodor, G. L. C. Apostol, J. Arabloo, M. Arafat, A. Y. Aravkin, D. Areda, A. Aremu, A. A. Artamonov, E. A. Ashley, M. O. Asika, S. S. Athari, M. M. W. Atout, T. Awoke, S. Azadnajafabad, J. M. Azam, S. Aziz, A. Y. Azzam, M. Babaei, F.-X. Babin, M. Badar, A. A. Baig, M. Bajcetic, S. Baker, M. Bardhan, H. J. Barqawi, Z. Basharat, A. Basiru, M. Bastard, S. Basu, N. S. Bayleyegn, M. A. Belete, O. O. Bello, A. Beloukas, J. A. Berkley, A. S. Bhagavathula, S. Bhaskar, S. S. Bhuyan, J. A. Bielicki, N. I. Briko, C. S. Brown, A. J. Browne, D. Buonsenso, Y. Bustanji, C. G. Carvalheiro, C. A. CastaÃśeda-Orjuela, M. Cenderadewi, J. Chadwick, S. Chakraborty, R. M. Chandika, S. Chandy, V. Chansamouth, V. K. Chattu, A. A. Chaudhary, P. R. Ching, H. Chopra, F. R. Chowdhury, D.-T. Chu, M. Chutiyami, N. Cruz-Martins, A. G. Da Silva, O. Dadras, X. Dai, S. D. Darcho, S. Das, F. P. De La Hoz, D. M. Dekker, K. Dhama, D. Diaz, B. F. R. Dickson, S. G. Djorie, M. Dodangeh, S. Dohare, K. G. Dokova, O. P. Doshi, R. K. Dowou, H. L. Dsouza, S. J. Dunachie, A. M. Dziedzic, T. Eckmanns, A. Ed-Dra, A. Eftekharimehrabad, T. C. Ekundayo, I. El Sayed, M. Elhadi, W. El-Huneidi, C. Elias,

S. J. Ellis, R. Elsheikh, I. Elsohaby, C. Eltaha, B. Eshrati, M. Eslami, D. W. Eyre, A. O. Fadaka, A. F. Fagbamigbe, A. Fahim, A. Fakhri-Demeshghieh, F. O. Fasina, M. M. Fasina, A. Fatehizadeh, N. A. Feasey, A. Feizkhah, G. Fekadu, F. Fischer, I. Fitriana, K. M. Forrest, C. Fortuna Rodrigues, J. E. Fuller, M. A. Gadanya, M. GajdÃącs, A. P. Gandhi, E. E. Garcia-Gallo, D. O. Garrett, R. K. Gautam, M. W. Gebregergis, M. Gebrehiwot, T. G. Gebremeskel, C. Geffers, L. Georgalis, R. M. Ghazy, M. Golechha, D. Golinelli, M. Gordon, S. Gulati, R. D. Gupta, S. Gupta, V. K. Gupta, A. D. Habtey-ohannes, S. Haller, H. Harapan, M. L. Harrison, A. I. Hasaballah, I. Hasan, R. S. Hasan, H. Hasani, A. H. Haselbeck, M. S. Hasnain, I. I. Hassan, S. Hassan, M. S. Hassan Zadeh Tabatabaei, K. Hayat, J. He, O. E. Hegazi, M. Heidari, K. Hezam, R. Holla, M. Holm, H. Hopkins, M. M. Hossain, M. Hosseinzadeh, S. Hostiuc, N. R. Hussein, L. D. Huy, E. D. IbÃąÃśez-Prada, A. Ikiroma, I. M. Ilic, S. M. S. Islam, F. Ismail, N. E. Ismail, C. D. Iwu, C. J. Iwu-Jaja, A. Jafarzadeh, F. Jaiteh, R. Jalilzadeh Yengejeh, R. D. G. Jamora, J. Javidnia, T. Jawaid, A. W. J. Jenney, H. J. Jeon, M. Jokar, N. Jomehzadeh, T. Joo, N. Joseph, Z. Kamal, K. K. Kanmodi, R. S. Kantar, J. A. Kapisi, I. M. Karaye, Y. S. Khader, H. Khajuria, N. Khalid, F. Khamesipour, A. Khan, M. J. Khan, M. T. Khan, V. Khanal, F. F. Khidri, J. Khubchandani, S. Khusuwan, M. S. Kim, A. Kisa, V. A. Korshunov, F. Krapp, R. Krumkamp, M. Kuddus, M. Kulimbet, D. Kumar, E. A. P. Kumaran, A. Kuttikkattu, H. H. Kyu, I. Landires, B. K. Lawal, T. T. T. Le, I. M. Lederer, M. Lee, S. W. Lee, A. Lepape, T. L. Lerango, V. S. Ligade, C. Lim, S. S. Lim, L. W. Limenh, C. Liu, X. Liu, X. Liu, M. J. Loftus, H. I. M Amin, K. L. Maass, S. B. Maharaj, M. A. Mahmoud, P. Maikanti-Charalampous, O. M. Makram, K. Malhotra, A. A. Malik, G. D. Mandilara, F. Marks, B. A. Martinez-Guerra, M. Martorell, H. Masoumi-Asl, A. G. Mathioudakis, J. May, T. A. McHugh, J. Meiring, H. N. Meles, A. Melese, E. B. Melese, G. Minervini, N. S. Mohamed, S. Mohammed, S. Mohan, A. H. Mokdad, L. Monasta, A. Moodi Ghalibaf, C. E. Moore, Y. Moradi, E. Mossialos, V. Mougin, G. D. Mukoro, F. Mulita, B. Muller-Pebody, E. Murillo-Zamora, S. Musa, P. Musicha, L. A. Musila, S. Muthupandian, A. J.

Nagarajan, P. Naghavi, F. Nainu, T. S. Nair, H. H. R. Najmuldeen, Z. S. Natto, J. Nauman, B. P. Nayak, G. T. Nchanji, P. Ndishimye, I. Negoi, R. I. Negoi, S. A. Nejadghaderi, Q. P. Nguyen, E. A. Noman, D. C. Nwakanma, S. O'Brien, T. J. Ochoa, I. A. Odetokun, O. A. Ogundijo, T. R. Ojo-Akosile, S. R. Okeke, O. C. Okonji, A. T. Olagunju, A. Olivas-Martinez, A. A. Olorukooba, P. Olwoch, K. I. Onyedibe, E. Ortiz-Brizuela, O. Osuolale, P. Ounchanum, O. T. Oyeyemi, M. P. P A, J. L. Paredes, R. R. Parikh, J. Patel, S. Patil, S. Pawar, A. Y. Peleg, P. Peprah, J. PerdigÃčo, C. Perrone, I.-R. Petcu, K. Phommasone, Z. Z. Piracha, D. Poddighe, A. J. Pollard, R. Poluru, A. Ponce-De-Leon, J. Puvvula, F. N. Qamar, N. H. Qasim, C. D. Rafai, P. Raghav, L. Rahbarnia, F. Rahim, V. Rahimi-Movaghar, M. Rahman, M. A. Rahman, H. Ramadan, S. K. Ramasamy, P. S. Ramesh, P. W. Ramteke, R. K. Rana, U. Rani, M.-M. Rashidi, D. Rathish, S. Rattanavong, S. Rawaf, E. M. M. Redwan, L. F. Reyes, T. Roberts, J. V. Robotham, V. D. Rosenthal, A. G. Ross, N. Roy, K. E. Rudd, C. J. Sabet, B. A. Saddik, M. R. Saeb, U. Saeed, S. Saeedi Moghaddam, W. Saengchan, M. Safaei, A. Saghazadeh, N. Saheb Sharif-Askari, A. Sahebkar, S. S. Sahoo, M. Sahu, M. Saki, N. Salam, Z. Saleem, M. A. Saleh, Y. L. Samodra, A. M. Samy, A. Saravanan, M. Satpathy, A. E. Schumacher, M. Sedighi, S. Seekaew, M. Shafie, P. A. Shah, S. Shahid, M. J. Shahwan, S. Shakoor, N. Shalev, M. A. Shamim, M. A. Shamshirgaran, A. Shamsi, A. Sharifan, R. P. Shastry, M. Shetty, A. Shittu, S. Shrestha, E. E. Siddig, T. Sideroglou, J. Sifuentes-Osornio, L. M. L. R. Silva, E. A. F. SimÃţes, A. J. H. Simpson, A. Singh, S. Singh, R. Sinto, S. S. M. Soliman, S. Soraneh, N. Stoesser, T. Z. Stoeva, C. K. Swain, L. Szarpak, S. S. T Y, S. Tabatabai, C. Tabche, Z. M.-A. Taha, K.-K. Tan, N. Tasak, N. Y. Tat, A. Thaiprakong, P. Thangaraju, C. C. Tigoi, K. Tiwari, M. R. Tovani-Palone, T. H. Tran, M. Tumurkhuu, P. Turner, A. J. Udoakang, A. Udoh, N. Ullah, S. Ullah, A. G. Vaithinathan, M. Valenti, T. Vos, H. T. L. Vu, Y. Waheed, A. S. Walker, J. L. Walson, T. Wangrangsimakul, K. G. Weerakoon, H. F. L. Wertheim, P. C. M. Williams, A. A. Wolde, T. M. Wozniak, F. Wu, Z. Wu, M. K. K. Yadav, S. Yaghoubi, Z. S. Yahaya, A. Yarahmadi, S. Yezli, Y. E. Yismaw, D. K. Yon, C.-W. Yuan,

H. Yusuf, F. Zakham, G. Zamagni, H. Zhang, Z.-J. Zhang, M. ZieliÅĎska, A. Zumla, S. H. H. Zyoud, S. H. Zyoud, S. I. Hay, A. Stergachis, B. Sartorius, B. S. Cooper, C. Dolecek, and C. J. L. Murray. Global burden of bacterial antimicrobial resistance 1990âĂŞ2021: a systematic analysis with forecasts to 2050. *The Lancet*, 404(10459): 1199–1226, Sept. 2024. ISSN 01406736. doi: 10.1016/S0140-6736(24)01867-1. URL https://linkinghub.elsevier.com/retrieve/pii/S0140673624018671.

M. Pelikan. *Hierarchical Bayesian optimization algorithm*. Springer, 2005.

R. Perumal and T. L. V. Zyl. Surrogate-assisted strategies: the parameterisation of an infectious disease agent-based model. *Neural Computing and Applications*, 37(2):627–638, Jan. 2025. ISSN 0941-0643, 1433-3058. doi: 10.1007/s00521-022-07476-y. URL https://link.springer.com/10.1007/s00521-022-07476-y.

G. Qian, M. Toizumi, S. Clifford, L. T. Le, T. Papastylianou, C. Satzke, B. Quilty, C. Iwasaki, N. Kitamura, M. Takegata, M. X. Bui, H. A. T. Nguyen, D. A. Dang, A. J. Van Hoek, L. M. Yoshida, and S. Flasche. Association of pneumococcal carriage in infants with the risk of carriage among their contacts in Nha Trang, Vietnam: A nested cross-sectional survey. *PLOS Medicine*, 19(5):e1004016, May 2022. ISSN 1549-1676. doi: 10.1371/journal.pmed.1004016. URL https://dx.plos.org/10.1371/journal.pmed.1004016.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. URL https://www.r-project.org/.

E. Ramos-Sevillano, G. Ercoli, and J. S. Brown. Mechanisms of Naturally Acquired Immunity to Streptococcus pneumoniae. *Frontiers in Immunology*, 10: 358, Mar. 2019. ISSN 1664-3224. doi: 10.3389/fimmu.2019.00358. URL https://www.frontiersin.org/article/10.3389/fimmu.2019.00358/full.

J.-H. Song, R. Dagan, K. P. Klugman, and B. Fritzell. The relationship between pneumococcal serotypes and antibiotic resistance. *Vaccine*, 30(17):2728–

2737, Apr. 2012. ISSN 0264410X. doi: 10.1016/j.vaccine.2012.01.091. URL https://linkinghub.elsevier.com/retrieve/pii/S0264410X12001648.

K. Van Zandvoort, A. I. Hassan, M. O. Bobe, C. L. Pell, M. S. Ahmed, B. D. Ortika, S. Ibrahim, M. I. Abdi, M. A. Karim, R. M. Eggo, S. Y. Ali, J. Hinds, S. M. Soleman, R. Cummings, C. R. McGowan, E. K. Mulholland, M. A. Hergeye, C. Satzke, F. Checchi, and S. Flasche. Pre-vaccination carriage prevalence of Streptococcus pneumoniae serotypes among internally displaced people in Somaliland: a cross-sectional study. *Pneumonia*, 16(1): 25, Dec. 2024. ISSN 2200-6133. doi: 10.1186/s41479-024-00148-6. URL https://pneumonia.biomedcentral.com/articles/10.1186/s41479-024-00148-6.

J. Vekemans, M. Hasso-Agopsowicz, G. Kang, W. P. Hausdorff, A. Fiore, E. Tayler, E. J. Klemm, R. Laxminarayan, P. Srikantiah, M. Friede, and M. Lipsitch. Leveraging Vaccines to Reduce Antibiotic Use and Prevent Antimicrobial Resistance: A World Health Organization Action Framework. *Clinical Infectious Diseases*, 73(4):e1011–e1017, Aug. 2021. ISSN 1058-4838, 1537-6591. doi: 10.1093/cid/ciab062. URL https://academic.oup.com/cid/article/73/4/e1011/6119610.

D. F. Vestrheim, E. A. HÃÿiby, I. S. Aaberge, and D. A. Caugant. Phenotypic and Genotypic Characterization of *Streptococcus pneumoniae* Strains Colonizing Children Attending Day-Care Centers in Norway. *Journal of Clinical Microbiology*, 46(8):2508–2518, Aug. 2008. ISSN 0095-1137, 1098-660X. doi: 10.1128/JCM.02296-07. URL https://journals.asm.org/doi/10.1128/JCM.02296-07.

E. Weerdenburg, T. Davies, B. Morrow, A. L. Zomer, P. Hermans, O. Go, B. Spiessens, T. Van Den Hoven, G. Van Geet, M. Aitabi, et al. Global distribution of o serotypes and antibiotic resistance in extraintestinal pathogenic escherichia coli collected from the blood of patients with bacteremia across multiple surveillance studies. *Clinical Infectious Diseases*, 76(3):e1236–e1243, 2023.