Life History Trade-offs in *Streptococcus pneumoniae*: Antibiotic Susceptibility and Vaccination

Martin Emons

Theoretical Biology Group
ETH Zürich
martin.emons@inf.ethz.ch

April 2021

1 Introduction

Streptococcus pneumoniae is a gram-positive bacterium that can colonise the upper respiratory tract asymptomatically. It can as well lead to infections with diseases such as bacterial pneumoniae, sepsis and meningitis. The most severe infections are collectively called invasive pneumococcal disease (IPD).[1, 2] Estimates on the worldwide death toll for children from the year 2009 range from 700'000 to 1'000'000 each year. [3] Especially high were the fatalities in children aged 1-59 months prior to vaccination with an estimated 8-12% of all death cases. IPD is as well a problem for people of older age with an estimated death toll of $\geq 50'000$ deaths in the US each year. [1, 2]

There are currently around 100 different serotypes known of *S. pneumoniae*. Because the manufacturing of conjugate vaccines is quite difficult, only a small fraction of all strains (7-13) are included in a vaccine.[2] This means that we can distinguish the strains of *S. pneumoniae* in terms of being included in the vaccine (vaccine-type strains) or not being included in the vaccine (non-vaccine-type strains).

A growing concern for public health is resistance in bacteria like *S. pneumoniae*. The economic costs induced by multi-drug resistance are estimated to be around $ 100 Trillion in the year 2050 in a worst case estimate. [4]. When trying to understand resistance one tries to understand the trade-offs resistance infers. On the one hand a resistant cell has a selective advantage through resistance, on the other hand it has a selective disadvantage through any costs the resistance incurs. Costs of resistance are believed to be fitness costs, meaning that they decrease the fitness of the strain, e.g. the per capita growth rate. [5]

This interplay between resistant and sensitive strains can be seen as a variant of the duration of infection vs. infectiousness trade-off. A sensitive strain has less fitness-costs since it does not have to maintain resistance leading to a higher infectiousness. Resistance on the other hand leads to an increased duration of infection since they are not sensitive to antibiotic treatment. [6, 7] Resistance comes at a competitive cost, meaning that the infectiousness of a resistant strain is decreased. In this trade-off (infectiousness or duration of infection) it is of course context dependent which process is more beneficial. [5]

One context in which the optimality of the trade-offs might change is the availability of hosts for the bacterium to colonise. Since vaccination changes the availability of hosts, it is a process of great interests for infectious disease dynamics. In 2000 the pneumococcal conjugate vaccine (PCV7) including 7 strains was admitted by the FDA. [8] This meant that suddenly the availability of hosts changed. For some types (vaccine-types) it decreased whereas for other types it increased (non-vaccine-types). This is what is called strain replacement. The positive effects due to the reduction of one strain are by-passed by the negative effects of strain replacement by those that are not included in the vaccine. [9]

In the context of strain replacement it is of interest to investigate the dynamics of antibiotic susceptibility. This leads to the main question of this report. How does the frequency in antibiotic susceptibility of *S. pneumoniae* change upon introduction of a vaccine? We will investigate this using an epidemiological model and try to verify our findings with data.

2 Results

2.1 SIS Model of Antibiotic Resistance

First, we analysed an SIS epidemiological model as introduced by Lehtinen et al. [4]. SIS stands for Susceptible-Infected-Susceptible. This model is suitable for modelling a lot of bacterial infections, including *S. pneumoniae* . An SIS model assumes that infected individuals are again susceptible after recovery and there is no immunity. Furthermore, there are no births or deaths, as these two forces are equalising. Next,the infected population can be subdivided into a compartment that is infected with resistant bacteria I_r and a compartment infected with susceptible bacteria I_s. The infections with sensitive bacteria can be treated with antibiotics at a rate τ, which increases their clearing rate. An assumption is that antibiotics lead to a direct clearance of the infection.

Since antibiotic resistance confers an advantage in clearance (can not use antibiotics) but a disadvantage in transmission (need to invest cost in resistance), we find that $\beta_s > \beta_r$. This is encoded via a cost term c_β resulting in $\beta_r = \frac{\beta}{c_\beta}$ and $\beta_s = \beta$. For the clearance rate we find that $\gamma_r > \gamma_s$. This is again moelled with a clearance cost c_γ and get $\gamma_r = c_\gamma \gamma$ and $\gamma_s = \gamma$. This leads to the following system of ordinary differential equations (ODEs) that describe the dynamics of antibiotic resistance as described by Lehtinen et al. [4].

$$\begin{aligned}
\frac{dS}{dt} &= (\gamma + \tau)I_s - \beta S I_s + c_\gamma \gamma I_r - \frac{\beta}{c_\beta} S I_r \\
\frac{dI_s}{dt} &= \beta S I_s - (\gamma + \tau) I_s \\
\frac{dI_r}{dt} &= \frac{\beta}{c_\beta} S I_r - c_\gamma \gamma I_r
\end{aligned} \tag{1}$$

The detailed description of the parameters can be found in table 1.

Parameter	Description
β	transmission rate
γ	clearance rate
τ	antibiotic consumption rate
c_β	cost of antibiotic resistance on transmission
c_γ	cost of antibiotic resistance on clearance

Table 1: The parameters of the SIS model are described in this table as a reference.

We can simplify this system by using the fact that we neglect births and deaths as seen above. This means we obtain a constant population size $N = S + I_r + I_s$. Thus, we end up with the following simplified system [4]:

$$\begin{aligned} \frac{dI_s}{dt} &= \beta(N - I_s - I_r)I_s - (\gamma + \tau)I_s \\ \frac{dI_r}{dt} &= \frac{\beta}{c_\beta}(N - I_s - I_r)I_r - c_\gamma \gamma I_r \end{aligned} \tag{2}$$

When we want to analyse the behaviour of an ODE system at equilibrium we want to analyse what values get, if the system does not change anymore, i.e. the time derivative equals to zero.

We obtain a disease free solution ($I_s = I_r = 0$) or three endemic solutions. The detailed derivation can be found in the appendix. The first two endemic solutions are out competition of either of the both strains, such that only one strain exists at equilibrium (appendix A). The interesting case is the endemic solutions that leads to coexistence.

$$S = \frac{\gamma + \tau}{\beta} = \frac{c_\gamma c_\beta \gamma}{\beta} \tag{3}$$

If we take the inverse of S it follows that $\frac{\beta}{\gamma+\tau} = \frac{\beta}{c_\gamma c_\beta \gamma}$. This is the definition of R_0 meaning that we get the relationship $R_{0_s} = R_{0_r}$ at equilibrium. This is an important point, since this result shows that if we want to obtain coexistence in the SIS system, the two R_0 values need to be the same. This fact will be needed for the numerical simulation of the SISV model [4].

This shows us all the possible combinations of equilibria that we could obtain and what values the variables would take dependent on the parameters. It can be seen that we have either out competition by sensitives or resistant strains if the initial conditions are larger than zero. The only case we can have coexistence, is when the R_0 values are the same for sensitive and resistant strains. For these equilibria we can determine the stability, which has been done in the appendix B. If $R_{0_r} > R_{0_s}$ the resistant type will outcompete the resistant type and vice versa [4].

Next to analysing the antibiotic model in equations 1 analytically we can look at the numerical solutions of the system. We integrated the system using the Runge-Kutta method and obtained the trajectories visible in figure 1. We see a drop in susceptibles early in the disease progression. In this phase we see a transient increase in sensitive infections and later on a domination by resistant infections. This already shows in very simple terms what we expect to happen. The initial dynamics are a sudden drop in availaibility of hosts. Due to this loss in susceptibles we see a transient increase of sensitive strains which are out competed by resistant strains in the long run.

We can as well perform a phase plane analysis of the simplified system described in equations 2. In figure 2 we see the trajectories of sensitive and resistant infected individuals for different starting conditions. The transient dynamics are highly dependent on the starting conditions on sensitive vs. resistant types. In the long run however we see the domination of the resistant strains no matter where we start from. This shows the importance of keeping the starting conditions in mind, as we are investigating transient dynamics.

Through the antibiotic resistance model in equation 1 and 2 we get an idea about how the sensitive and resistant strains behave subject to a decrease in available host. Via the phase plane analysis in figure 2 we understand that initial conditions matter a lot for transient phenomena. The flaw in the modelling up until now is, that we do not account for vaccination explicitly. This leads to the next, more complex SISV model.

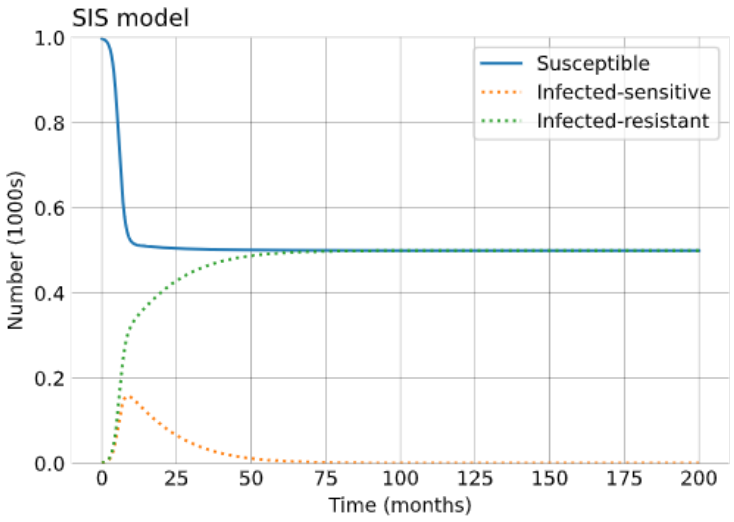

Figure 1: The trajectories of the SIS ODE system are plotted. On the y-axis are the number of individuals per compartment in 1000s. The x-axis is the time measured in months. We see a sharp drop in susceptibles in the first 10 months and an increase in both infected compartments. After approx. 10 months the infected-sensitive cases go back again and the infected-resistant cases make up 50% at equilibrium. The parameter values are $c_\beta = 1.0$, $c_\gamma = 1.0$, $\beta = 2/N$, $\gamma = 1$ and $\tau = 0.075$

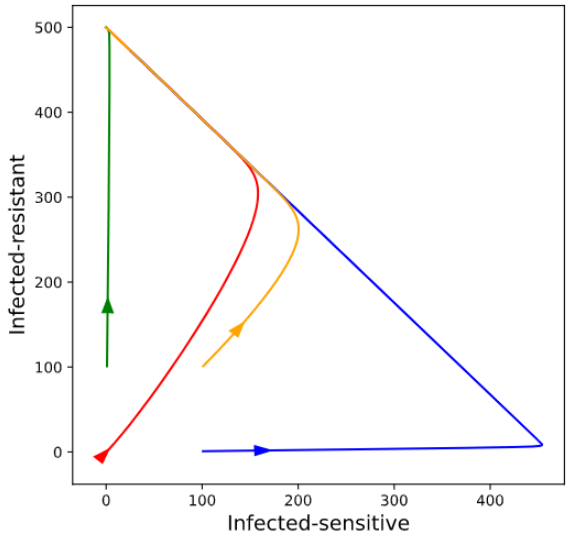

Figure 2: This graph shows a phase plane analysis of the simplified system described in equations 2. According to different starting conditions (the points in the lower-left quadrant where the trajectories start) we see different transient dynamics. In the long run we see however, that the resistant types will out compete the sensitives

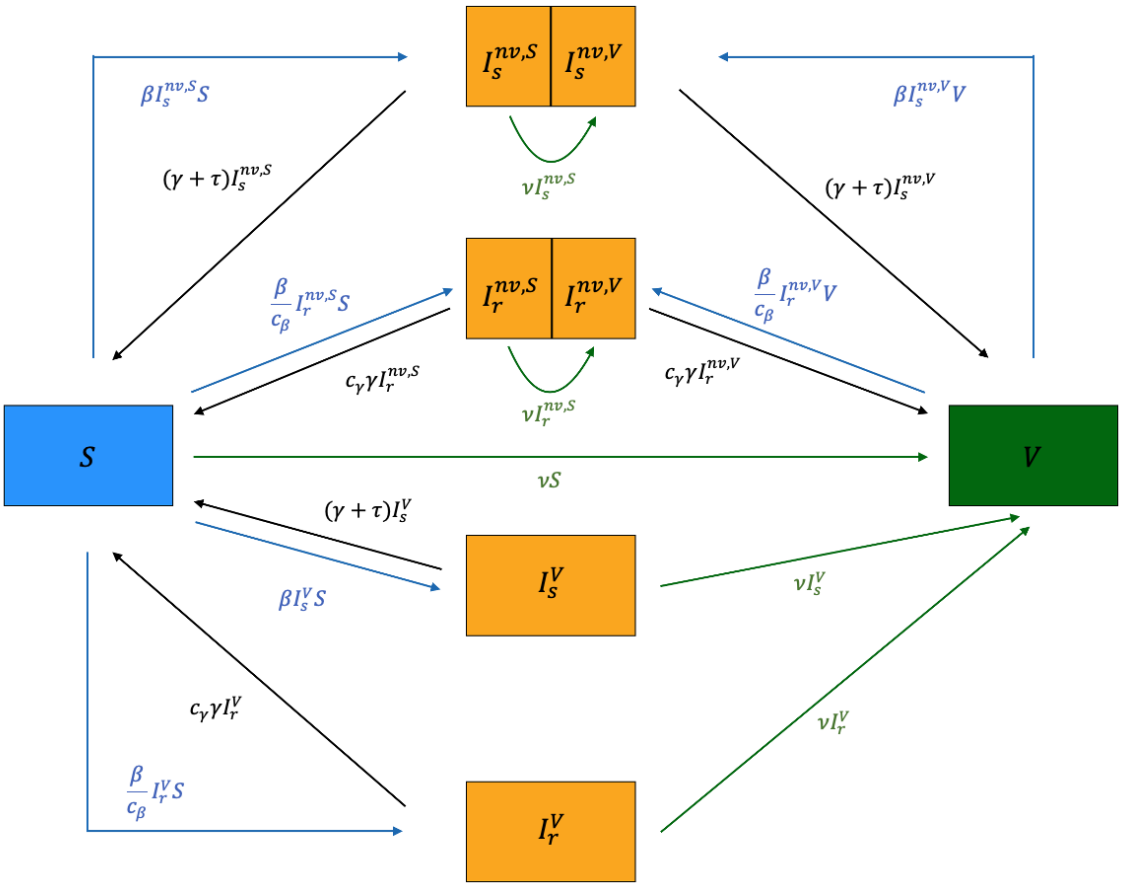

Figure 3: Flowchart describing the equations from the system 4. The variables are described in table 2 and the parameters in table 3. The equations of the ODE system are defined in equation 4.

2.2 SISV Model of Antibiotic Resistance and of Vaccination

The main idea of this model was to take the SIS model from equation 1 and allow for vaccination. This gives rise to an SISV model, which includes a vaccinated compartment V. Since a vaccine against *S. pneumoniae* only includes certain strains and not others, we had to distinguish between infections with a vaccine-type I^v (meaning a strain that is included in the vaccine) and a non-vaccine-type I^{nv}(not included in the vaccine). The dynamics of resistant and sensitive strains happens in the same way for non-vaccine-type and vaccine-type as described in the SIS model (equation 1). Vaccinating individuals happens at a rate ν. Individuals that were vaccinated can not be infected again with a vaccine-type bacterium. In order to keep track of the non-vaccine-type infections, a subcompartment was created e.g. $I^{nv,S}$ and $I^{nv,V}$ for non-vaccinated individual with infections and vaccinated individuals with infections respectively. This gave rise to the following system of ODEs described in the flowchart 3 and in the equations of the equation 4.

$$
\begin{aligned}
\frac{dS}{dt} &= (\gamma+\tau)(I_s^{nv,S}+I_s^v) - \beta S(I_s^{nv,S}+I_s^v) + c_\gamma\gamma(I_r^{nv,S}+I_r^v) - \frac{\beta}{c_\beta}S(I_r^{nv,S}+I_r^v) - \nu S \\
\frac{dI_s^{nv,S}}{dt} &= \beta S I_s^{nv,S} - (\gamma+\tau)I_s^{nv,S} - \nu I_s^{nv,S} \\
\frac{dI_r^{nv,S}}{dt} &= \frac{\beta}{c_\beta} S I_r^{nv,S} - c_\gamma\gamma I_r^{nv,S} - \nu I_r^{nv,S} \\
\frac{dI_s^{nv,V}}{dt} &= \beta V I_s^{nv,V} - (\gamma+\tau)I_s^{nv,V} + \nu I_s^{nv,S} \\
\frac{dI_r^{nv,V}}{dt} &= \frac{\beta}{c_\beta} V I_r^{nv,V} - c_\gamma\gamma I_r^{nv,V} + \nu I_r^{nv,S} \\
\frac{dI_s^{v}}{dt} &= \beta S I_s^{v} - (\gamma+\tau)I_s^{v} - \nu I_s^{v} \\
\frac{dI_r^{v}}{dt} &= \frac{\beta}{c_\beta} S I_r^{v} - c_\gamma\gamma I_r^{v} - \nu I_r^{v} \\
\frac{dV}{dt} &= \nu S + (\gamma+\tau)I_s^{nv,V} - \beta V I_s^{nv,V} + c_\gamma\gamma I_r^{nv,V} - \frac{\beta}{c_\beta} V I_r^{nv,V} + \nu(I_s^v + I_r^v)
\end{aligned}
\tag{4}
$$

The variables of equation 4 are described in table 2 and the parameters in table 3.

Variable	Description
S	Susceptibles
$I_s^{nv,S}$	Unvaccinated, infected with a sensitive strain of the non-vaccine-type
$I_r^{nv,S}$	Unvaccinated, infected with a resistant strain of the non-vaccine-type
$I_s^{nv,V}$	Vaccinated, infected with a sensitive strain of the non-vaccine-type
$I_r^{nv,V}$	Vaccinated, infected with a resistant strain of the non-vaccine-type
I_s^v	Infected with a resistant strain of the vaccine-type
I_r^v	Infected with a sensitive strain of the vaccine-type
V	Vaccinated

Table 2: The variables of the SISV model are described in this table as a reference.

Parameter	Description
β	transmission rate
γ	clearance rate
τ	antibiotic consumption rate
c_β	cost of antibiotic resistance on transmission
c_γ	cost of antibiotic resistance on clearance
ν	vaccination rate

Table 3: The parameters of the SISV model are described in this table as a reference.

If we want to introduce a rollout of the vaccine after reaching equilibrium, we need to have sensitive and resistant individuals. In order to do this we will enforce coexistence. We saw that the R_0 values for the sensitives and the resistants need to be the same as we have derived in equation 3: $\frac{\beta}{\gamma+\tau} = \frac{\beta}{c_\gamma c_\beta \gamma}$. We obtain the relationship $R_{0_s} = R_{0_r}$ at equilibrium and in this state we will reach the equilibrium of coexistence.

This leads to a constraint on the antibiotic consumption rate as done by Lehtinen et al. [4]:

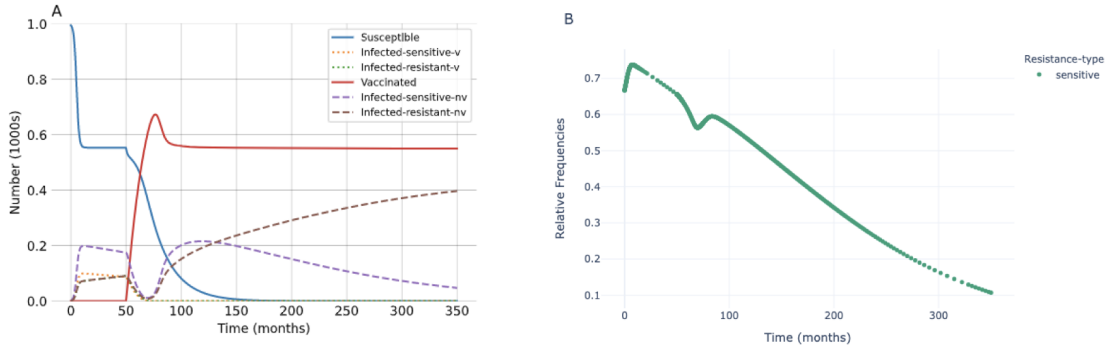

Figure 4: (A)Plot of the dynamics of SISV model over time. Vaccination after timepoint $t = 50$. On the y-axis are the number of individuals per compartment in 1000s. The x-axis is the time measured in months. The parameter values are $c_\beta = 1.1$, $c_\gamma = 1.0$, $\beta = 2/N$, $\gamma = 1$, $\tau = 0.11$ and $\nu = 0.05$. In this plot the subcompartments are added together, so that the behaviour of vaccine-types vs non-vaccine-types in terms of susceptibility can be analysed more easily. The initial conditions were $I_s^{nv,S} = 2, I_r^{nv,S} = 1, I_s^{nv,V} = 0, I_r^{nv,V} = 0, I_s^v = 1, I_r^v = 1$ and $V = 0$. (B) Plot of the relative frequency of sensitive types in the non-vaccine compartment. The x-axis is the time in months and the y-axis the frequency of sensitives relative to the total amount of infected individuals(sensitives and resistant strains)

$$\begin{aligned} R_{0_s} &= R_{0_r} \\ \frac{\beta}{\gamma + \tau} &= \frac{\beta}{c_\gamma c_\beta \gamma} \\ \tau &= \gamma(c_\gamma c_\beta - 1) \end{aligned} \tag{5}$$

In the beginning the two vaccinated and susceptible Infected subcompartments evolve in equilibrium. This means the parameter τ is set to be as above and ν to be zero so that there are no vaccinated indivudals. In order not to have too much confounding by the coexistence assumption, a very small competition in the value of $\Delta = 0.01$ is introduced which leads to the strains still coexisting but in the long term after vaccination competitive effects are visible.

This system can again be analysed numerically and we obtain the following trajectories visible in figure 4. In the first 50 months the trajectories are coexisting with the infected sensitives individuals that were not vaccinating have a little advantage, du to the doubled initial condition. This leads to $\sim 2x$ sensitive to resistant non-vaccine types when the vaccination starts. The introduction of the vaccination campaign at $t = 50$ shows a drop in susceptibles and a rise in vaccinated individuals. Interestingly, the vaccine-type strains decline first before the non-vaccine-type strains take over the field. This period of change in availability of host shows a transient dominance of sensitives over resistant strains in the non-vaccine-type strains. In the long run the resistant strains will be fixed however.

The simulation gives rise to the hypothesis that a transient shortage of host, as is the case during an ongoing vaccination campaign, will lead to a transient increase of sensitive frequencies an a slower increase in resistant strains. In order to verify this claim, we looked at data from pneumococcal vaccination campaigns.

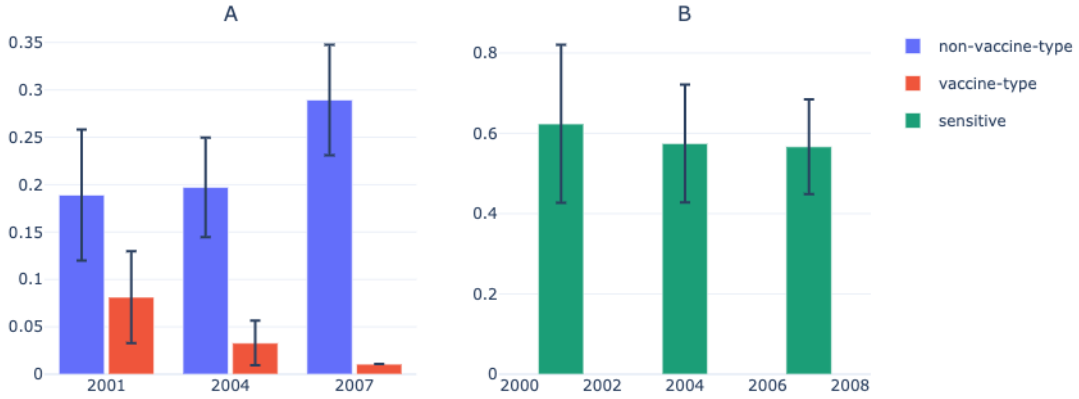

Figure 5: (A) Plotted here are the relative frequencies of vaccine-types vs. non-vaccine-types over the course of the collection period. The relative frequency was corrected for different test positivity over the years (27%, 23% and 30%).(B) Plotted here are the relative frequencies of resistant vs. sensitive strains over the course of the collection period. The Binomial Confidence interval is indicated to show the uncertainty in the data.

2.3 Data Analysis

First we look at the development of vaccine-types vs. non-vaccine-types over the course of time. This is visible in figure 5 (A). The dataset is from Massachusetts hospitals, collected in the form of nasal swaps in children. The vaccination campaign with PCV7 started in 2000. Therefore, we make the assumption that the year 2001 constitutes as a pre-vaccination data point. Thus, the pre-vaccination frequencies are such, that the non-vaccine-types are already more common than the vaccine-types. This is something we can as well enforce in the simulation in figure 4 by changing the initial conditions. It is visible that first the vaccine-type frequencies decline before the non-vaccine-type frequencies rise. This observation was as well seen in the simulation in figure 4.

The interest in the transition for the non-vaccine-types as seen in the simulation (figure 4) meant that in the following analysis we focused on non-vaccine-types. This was due to the hypothesis on the transient domination of sensitive strains in the non-vaccine types. The overall amount of positive *S. pneumoniae* cases increased over the years, which was accounted for in figure 5 (A). In order to compare the years, the relative frequencies of sensitive strains in the non-vaccine-type were computed in figure 5 (B). In 2001 there are already more sensitive than resistant strains. There is a trend of equilibration, but this trend is not statistically significant. A chi-squared test on the absolute counts gave a p-value of 0.62.

3 Discussion

Competitive exclusion and transient increase in sensitives Unless we make the simplifying assumption of having the exact same R_0 values for the resistant and the sensitive strains, we expect a competitive exclusion. Depending on which R_0 value is larger we get either an exclusion of

sensitives or resistant strains. This can be seen in figure 1. Given the initial conditions the parameter values we see a competitive exclusion of the sensitive strain. The interesting thing is the transient behaviour. There, we see a peak in sensitives which decreases again after some time. We hypothesise that this is due to the shortage in susceptibles. The drop of susceptibles leads to a transient increase in sensitives but in the long run the resistant strains out compete the sentitives. This process is analogous to vaccination which leads as well to a shortage in host. One thing we would have expected to see was a transient dominance, meaning that the sensitive grow faster than the resistant strains. When checking for this, we that I_s only grows faster than I_r if the inequality $\beta S_{df}\left(1-\frac{1}{c_\beta}\right)+\gamma(c_\gamma-1)-\tau>0$ holds (see Appendix C). This was not given by our parameter choices.

Initial conditions matter for transient peak In figure 2 the trajectories for different initial conditions are plotted. It is clear the transient behaviour is highly dependent on the initial conditions. This is of importance because we can adjust this to our data, since we have as well a pre-vaccination datapoint, so we can integrate the initial conditions into our model.

Vaccine model gives transient dominance of sensitives Using the model of vaccination and antibiotic resistance in equation 4 and initialising with twice as many sensitive as resistant strains (as found in the data in figure 5 (B)) we observe a transient dominance of sensitives in figure 4. In the long term we parameterised the model such that we see competitive exclusion of sensitives, as this is what we expect to happen in nature. Furthermore, we see the decline in the vaccine type strains prior to the increase in the non-vaccine types.

Prior Decline in vaccine-type found in data The decline in vaccine-type as found in the simulation figure 4 can be found in the data as well in figure 5 (A). From 2001 to 2004 the vaccine-types decline whereas the non-vaccine-types remain more or less constant. The non-vaccine-types only start to increase substantially in the interval 2004-2007. Overall, the vaccine types go down after introduction of the vaccine, whereas the non-vaccine types increase. In more detail we see that the vaccine-types decrease prior to the increase in the non-vaccine types. This finding supports thus the claim in our simulation.

Relative frequencies are comparable in simulation and in data In figure 5 (B)

Transient dominance of sensitives is not found in the data The transient dominance as found in the simulation figure 4 is not found in the data figure 5 (B). We see that in the beginning already we have more sensitive strains than resistant ones (which we accounted for in the simulation with our initial conditions). Then we see though no increase in sensitives but rather a decrease and an increase in resistant types. This is not aligned with our hypothesis. There are multiple ways to explain this mismatch. Firstly, the timespan between the data points is quite large (4 year intervals or 48 months). In the simulation the vaccination campaign starts at $t = 50$ and the plateau after which we expect a decline in sensitives again is at $\sim t = 100$. This means in the simulation, we expect the transient behaviour to show within approximately 50 months. Since the data was only collected 48 months after the first timepoint, it could be that we simply miss the transient phase. This should be investigated in more detail with a more fine-grained dataset, where we ideally would have a datapoint each year. The other explanaition might be, that our model in equation 4 was too simplified, since we had introduce coexistence by setting the treatment rate to equation 3. This would mean that one could set up a more sophisticated model taking into account aspects such as balancing selection [4] to explain coexistence, which we did not do in the course of this project.

Parameter choices In the course of this discussion we have made claims that are dependent on the choice of parameters. Thus, we want to discuss how sensible the choice of parameters is in the context of *S. pneumoniae*. The values for β, τ and γ were taken from Colijn et al. [10] and Lehtinen et al. [4]. Colijn et al. describe in detail which parameter values make sense in which context. The cost parameters c_β and c_γ were chosen accordingly to Lehtinen et al. [4]. The choices of the costs are though important as certain behaviours can only be seen for certain values of c_β and c_γ (Appendix figure 6).

4 Conclusion

In this project we have seen an approach to modelling the effect vaccination has on antibiotic resistance. This model gave rise to the hypothesis that after vaccination we should see a transient dominance of sensitive strains followed by an out competition of those strains when going towards equilibrium. This finding is not found in our dataset. Furthermore, we hypothesise from the simulation that the vaccine-types should decrease prior to the non-vaccine-types increasing. This can be found in our data analysis. Further work should focus on data analysis of a more fine-grained set and work on a more complex model including different approaches to achieve coexistence.

5 Acknowledgments

I thank Sonja Lehtinen for the kind supervision, it has been a very pleasant working atmosphere and a stimulating research topic. I thank as well Sebastian Bonhoeffer for making the rotation possible in the theoretical biology group.

6 Material & Methods

6.1 Datasets and Data Availability

The data is made available on the github page of the project `https://github.com/mjemons/IDD_TB` in the subfolder `data`. The data used was the Massachusetts dataset described in Croucher et al.[11]. It contains information on nasal swaps of 616 children collected between 2001 and 2007.

6.2 Model Implementation and Code Availability

The models and data analysis was implemented using python 3.7.9. All the packages and their corresponding versions are documented in the `requirements.txt` on the following github page `https://github.com/mjemons/IDD_TB` in the subfolder `scripts`. All code that was used in this project can be found on the same github page.

6.3 Parameterisation

The model parameters chosen were taken from Lehtinen et al. [4] and Colijn et. al. [10]. They are given as rates per month (see as well discussion). The initial conditions were chosen to be always 1 for the infecteds except for the SISV numerical integration. There the $I_s^{nv,S}$ type was set to 2, in order to get conditions at $t = 50$ that are comparable with the data when the vaccination campaign started.

References

[1] Paul N. Zivich, John D. Grabenstein, Sylvia I. Becker-Dreps, and David J. Weber. Streptococcus pneumoniae outbreaks and implications for transmission and control: a systematic review. *Pneumonia*, 10(1), 2018.

[2] Alessandra Løchen, Nicholas J. Croucher, and Roy M. Anderson. Divergent serotype replacement trends and increasing diversity in pneumococcal disease in high income settings reduce the benefit of expanding vaccine valency. *Scientific Reports*, 10(1):1–17, 2020.

[3] Katherine L. O'Brien, Lara J. Wolfson, James P. Watt, Emily Henkle, Maria Deloria-Knoll, Natalie McCall, Ellen Lee, Kim Mulholland, Orin S. Levine, and Thomas Cherian. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *The Lancet*, 374(9693):893–902, 2009.

[4] Sonja Lehtinen, François Blanquart, Nicholas J. Croucher, Paul Turner, Marc Lipsitch, and Christophe Fraser. Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proceedings of the National Academy of Sciences of the United States of America*, 114(5):1075–1080, 2017.

[5] Andrew D. Letten, Alex R. Hall, and Jonathan M. Levine. Using ecological coexistence theory to understand antibiotic resistance and microbial competition. *Nature Ecology and Evolution*, 5(April):431–441, 2021.

[6] Sergey V. Melnikov, David L. Stevens, Xian Fu, Hui Si Kwok, Jin Tao Zhang, Yue Shen, Jeffery Sabina, Kevin Lee, Harry Lee, and Dieter Söll. Exploiting evolutionary trade-offs for posttreatment management of drug-resistant populations. *Proceedings of the National Academy of Sciences of the United States of America*, 117(30):17924–17931, 2020.

[7] Christophe Fraser, Katrina Lythgoe, Gabriel E. Leventhal, George Shirreff, T. Déirdre Hollingsworth, Samuel Alizon, and Sebastian Bonhoeffer. Virulence and pathogenesis of HIV-1 infection: An evolutionary perspective. *Science*, 343(6177), 2014.

[8] Centers for Disease Control and Prevention. About Pneumococcal Vaccines.

[9] Yonas I. Tekle, Kaare M. Nielsen, Jingzhou Liu, Melinda M. Pettigrew, Lauren A. Meyers, Alison P. Galvani, and Jeffrey P. Townsend. Controlling Antimicrobial Resistance through Targeted, Vaccine-Induced Replacement of Strains. *PLoS ONE*, 7(12):1–9, 2012.

[10] Caroline Colijn, Ted Cohen, Christophe Fraser, William Hanage, Edward Goldstein, Noga Givon-Lavi, Ron Dagan, and Marc 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, 2010.

[11] Nicholas J. Croucher, Jonathan A. Finkelstein, Stephen I. Pelton, Julian Parkhill, Stephen D. Bentley, Marc Lipsitch, and William P. Hanage. Population genomic datasets describing the post-vaccine evolutionary epidemiology of Streptococcus pneumoniae. *Scientific Data*, 2:1–9, 2015.

A Derivation of SIS Equilibria

Since we neglect births and deaths as seen above we get a constant population size N: $N = S+I_r+I_s$. This means the system becomes a 2x2 system of equations:

$$\frac{dI_s}{dt} = \beta(N - I_s - I_r)I_s - (\gamma + \tau)I_s$$
$$\frac{dI_r}{dt} = \frac{\beta}{c_\beta}(N - I_s - I_r)I_r - c_\gamma\gamma I_r$$

Factoring out I_r and I_s we get:

$$\frac{dI_s}{dt} = I_s\left(\beta(N - I_s - I_r) - (\gamma + \tau)\right) \stackrel{!}{=} 0$$
$$\frac{dI_r}{dt} = I_r\left(\frac{\beta}{c_\beta}(N - I_s - I_r) - c_\gamma\gamma\right) \stackrel{!}{=} 0$$

Here we get either a disease free solution ($I_s = I_r = 0$) or three endemic solutions

- disease free: $I_s = I_r = 0$
- endemic I: $I_s = I_r \neq 0$

$$N - I_s - I_r = \frac{\gamma + \tau}{\beta}$$
$$N - I_s - I_r = \frac{c_\gamma c_\beta \gamma}{\beta}$$

It follows that $S = \frac{\gamma+\tau}{\beta} = \frac{c_\gamma c_\beta \gamma}{\beta}$

If we take the inverse of S it follows that $\frac{\beta}{\gamma+\tau} = \frac{\beta}{c_\gamma c_\beta \gamma}$. This is the definition of R_0 meaning that we get the relationship $R_{0_s} = R_{0_r}$ at equilibrium.

- endemic II: $I_s = 0, I_r \neq 0$

$$\frac{dI_s}{dt} = I_s\left(\beta(N - I_s - I_r) - (\gamma + \tau)\right) \stackrel{!}{=} 0$$
$$\frac{dI_r}{dt} = I_r\left(\frac{\beta}{c_\beta}(N - I_s - I_r) - c_\gamma\gamma\right) \stackrel{!}{=} 0$$

Since $I_s = 0$ we get:

$$I_r\left(\frac{\beta}{c_\beta}(N - I_r) - c_\gamma\gamma\right) = 0$$

since $I_r \neq 0$ the brackets need to zero:

$$\frac{\beta}{c_\beta}(N - I_r) - c_\gamma\gamma = 0$$
$$\frac{\beta}{c_\beta}N - \frac{\beta}{c_\beta}I_r - c_\gamma\gamma = 0$$
$$\frac{\beta}{c_\beta}I_r = \frac{\beta}{c_\beta}N - c_\gamma\gamma$$
$$I_r = N - \frac{c_\beta c_\gamma \gamma}{\beta}$$

- endemic III: $I_s \neq 0, I_r = 0$

$$\frac{dI_S}{dt} = I_s \left(\beta(N - I_s - I_r) - (\gamma + \tau)\right) \stackrel{!}{=} 0$$
$$\frac{dI_R}{dt} = I_r \left(\frac{\beta}{c_\beta}(N - I_s - I_r) - c_\gamma \gamma\right) \stackrel{!}{=} 0$$

Since $I_r = 0$ we get:

$$I_s \left(\beta(N - I_s) - (\gamma + \tau)\right) = 0$$

since $I_s \neq 0$ the brackets need to zero:

$$\begin{aligned} \beta(N - I_s) - (\gamma + \tau) &= 0 \\ \beta N - \beta I_s - \gamma + \tau &= 0 \\ \beta I_s &= \beta N - \gamma + \tau \\ I_s &= N - \frac{\gamma + \tau}{\beta} \end{aligned}$$

B Derivation of SIS Stability

Stability analysis will allow us to decide which equilibrium will be attained according to the parameters. In this analysis we perform a linearisation of the system around the equilibrium point. First, we will define the Jacobian Matrix

$$\begin{aligned} J &= \begin{pmatrix} \frac{\partial f_1}{\partial I_s} & \frac{\partial f_1}{\partial I_r} \\ \frac{\partial f_2}{\partial I_s} & \frac{\partial f_2}{\partial I_r} \end{pmatrix} \\ &= \begin{pmatrix} \beta(N - 2I_s - I_r) - (\gamma + \tau) & -\beta I_s \\ -\frac{\beta}{c_\beta} I_r & \frac{\beta}{c_\beta}(N - I_s - 2I_r) - c_\gamma \gamma \end{pmatrix} \end{aligned}$$

the stability of the equilibrium point can be determined by the sign of the eigenvalues of the Jacobian. In order to solve this, we will have to define the characteristic polynomial:

$$\det(J - \lambda I) = 0$$

In order to solve this we will plug in the equilibrium solutions and then solving the resulting characteristic polynomial.

We will analyse it first generally and then plug in the corresponding values at the equilibria.

$$\begin{aligned} &\det(J - \lambda I) = 0 \\ &= \begin{pmatrix} \beta(N - 2I_s - I_r) - (\gamma + \tau) - \lambda & -\beta I_s \\ -\frac{\beta}{c_\beta} I_r & \frac{\beta}{c_\beta}(N - I_s - 2I_r) - c_\gamma \gamma - \lambda \end{pmatrix} = 0 \\ &= \left[\beta(N - 2I_s - I_r) - (\gamma + \tau) - \lambda\right] \cdot \left[\frac{\beta}{c_\beta}(N - I_s - 2I_r) - c_\gamma \gamma - \lambda\right] - \left[-\beta I_s\right] \cdot \left[-\frac{\beta}{c_\beta} I_r\right] = 0 \\ &= \lambda^2 + \lambda \left(c_\gamma \gamma - \frac{\beta}{c_\beta}(N - I_s - 2I_r) + (\gamma + \tau) - \beta(N - 2I_s - I_r)\right) + \frac{\beta^2}{c_\beta}(N - 2I_s - I_r)(N - I_s - 2I_r) \\ &- \beta c_\gamma \gamma (N - 2I_s - I_r) + c_\gamma \gamma (\gamma + \tau) - (\gamma + \tau)\frac{\beta}{c_\beta}(N - I_s - 2I_r) - \frac{\beta^2}{c_\beta} I_r I_s \end{aligned}$$

this can then be solved using midnights formula:

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The parameters a, b, c for the different equilibrium points are:
* disease free ($I_s = I_r = 0$) :

$$\begin{aligned} a &= 1 \\ b &= \frac{\beta}{c_\beta} N + c_\gamma \gamma + \gamma + \tau - \beta N \\ c &= \frac{\beta^2}{c_\beta} N^2 - \beta c_\gamma \gamma N - \frac{\beta}{c_\beta} \gamma N + c_\gamma(\gamma + \tau) - \frac{\beta}{c_\beta} N(\tau + \gamma) \end{aligned}$$

- endemic I ($I_s \neq 0, I_r \neq 0$):

$$\begin{aligned} a &= 1 \\ b &= c_\gamma \gamma - \frac{\beta}{c_\beta}(N - I_s - 2I_r) + (\gamma + \tau) - \beta(N - 2I_s - I_r) \\ c &= \frac{\beta^2}{c_\beta}(N - 2I_s - I_r)(N - I_s - 2I_r) - \beta c_\gamma \gamma (N - 2I_s - I_r) \\ &\quad + c_\gamma \gamma(\gamma + \tau) - (\gamma + \tau)\frac{\beta}{c_\beta}(N - I_s - 2I_r) - \frac{\beta^2}{c_\beta} I_r I_s \end{aligned}$$

- endemic II ($I_s = 0, I_r \neq 0$):

$$\begin{aligned} a &= 1 \\ b &= c_\gamma \gamma - \frac{\beta}{c_\beta}(N - 2I_r) + (\gamma + \tau) - \beta(N - I_r) \\ c &= \frac{\beta^2}{c_\beta}(N - I_r)(N - 2I_r) - \beta c_\gamma \gamma (N - I_r) \\ &\quad + c_\gamma \gamma(\gamma + \tau) - (\gamma + \tau)\frac{\beta}{c_\beta}(N - 2I_r) - \frac{\beta^2}{c_\beta} \end{aligned}$$

- endemic III ($I_s \neq 0, I_r = 0$):

$$\begin{aligned} a &= 1 \\ b &= c_\gamma \gamma - \frac{\beta}{c_\beta}(N - I_s) + (\gamma + \tau) - \beta(N - 2I_s) \\ c &= \frac{\beta^2}{c_\beta}(N - 2I_s)(N - I_s) - \beta c_\gamma \gamma (N - 2I_s) \\ &\quad + c_\gamma \gamma(\gamma + \tau) - (\gamma + \tau)\frac{\beta}{c_\beta}(N - I_s) \end{aligned}$$

C Initial Growth analysis

We can analyse the initial growth of the system as follows

$$\frac{dS}{dt} = (\gamma + \tau)I_s - \beta S I_s + c_\gamma \gamma I_r - \frac{\beta}{c_\beta} S I_r$$
$$\frac{dI_s}{dt} = \beta S I_s - (\gamma + \tau) I_s$$
$$\frac{dI_r}{dt} = \frac{\beta}{c_\beta} S I_r - c_\gamma \gamma I_r$$

We plug in the initial conditions of the disease free state
if I_s should grow faster than I_r we get the following

$$\beta S_{df} - (\gamma + \tau) > \frac{\beta}{c_\beta} S_{df} - c_\gamma \gamma$$

as $S_{df} = C$ (the initial condition of S) in the disease-free state we get

$$\beta S_{df} - \frac{\beta}{c_\beta} S_{df} - \gamma - \tau + c_\gamma \gamma > 0$$
$$\beta S_{df} \left(1 - \frac{1}{c_\beta}\right) + \gamma(c_\gamma - 1) - \tau > 0$$

This means we only suspect a faster growth of I_s in the beginning (dominance) if the above holds true. This is the case e.g. when $c_\gamma = c_\beta \geq 1$. If only one of them is greater than 1, then the inequality above has to hold.

D Sensitivity Analysis of Parameters

Figure 6: Sensitivity analysis for different cost values. The Riemann sum of the sensitive curve was computed or different values of c_γ and c_β.