Department of Mathematics
and Statistics,
University of Strathclyde

Immunisation, Hepatitis and
Blood Safety Department,
Health Protection
Agency

PhD Thesis

Modelling the ecology and epidemiology of *Streptococcus pneumoniae* and their impact on conjugate vaccination

Author:
Stefan Flasche

Supervisors:
Yoon H. Choi
W. John Edmunds
Chris Robertson

July 30, 2012

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Stefan Flasche

July 30, 2012

Acknowledgements

I am grateful to my supervisors Dr. Yoon H. Choi, Prof. Chris Robertson and Prof W. John Edmunds for their scientific advice, for their encouragement and for them supporting my viral detour.

I would like to add a special thanks to Liz Miller for all the hours of inspiring discussions and for allowing me an unparalleled insight into public health surveillance and decision making.

Thanks to my colleagues at the Health Protection Agency. You were of magnificent support and provided a great working environment. I enjoyed my time at HPA a lot to which you contributed a great deal. Thanks to both AJ and Marc for plenty of helpful discussions and lots of highly appreciated distraction.

Ein grosses Danke an meine Familie und insbesondere an meine Mom, die mich durchwegs in meinen Vorhaben bestaerkt hat und die Anfertigung dieser Arbeit ermoeglicht hat. Dank auch an meine Berliner Freunde, die mir trotz der Entfernung immer nahe waren und die immer ein offenes Ohr und ne Flasche Ramazotti fuer mich hatten.

Meinem Mausi danke ich, dass wir dieses Abenteuer gemeinsam angegangen sind. Deine Unterstuetzung ist einzigartig!

Contents

List of Figures

List of Tables

Previously published work

The work presented in this thesis is the result of original research carried out by the author, Stefan Flasche, unless stated otherwise.

Supervision

The research was carried out under the supervision of Dr. Yoon Hong Choi from the Health Protection Agency, Professor William John Edmunds from the London School of Hygiene and Tropical Medicine and Health Protection Agency and Professor Chris Robertson from the University of Strathclyde and Health Protection Scotland.

Publications

Publications arising from this thesis are listed below. This work usually included the contribution of several people and the role of S. Flasche is outlined below. Additional publications arose from work on pandemic H1N1 influenza in 2009. Although carried out during the time of the PhD this work is not included in the thesis as it was conducted mostly during half a year of additional funding from the Health Protection Agency.

Yoon H. Choi, Marc Jit, Stefan Flasche, Nigel Gay and Elizabeth Miller. Mathematical modelling long-term effects of replacing 7-valent pneumococcal conjugate vaccine, Prevnar 7 with Prevnar 13 on invasive pneumococcal diseases in the England and Wales. *PLoS ONE*, 7:e39927, July 2012
S. Flasche transformed the deterministic model into an equivalent individual based model and assisted with the writing of the manuscript. In parts presented in section 5.3.

Stefan Flasche, W. John Edmunds, Elizabeth Miller, David Goldblatt, Chris Robertson and Yoon H. Choi. The impact of specific and non-specific immunity on the ecology of S. pneumoniae and the implications for vaccination. (submitted)
S. Flasche designed and implemented the model, carried out the analyses and wrote the manuscript. Presented in detail in sections 6.1 to 6.4.

Stefan Flasche, Mary Slack, and Elizabeth Miller. Long term trends introduce a potential bias when evaluating the impact of the pneumococcal conjugate vaccination programme in England and Wales. *Eurosurveillance*, 16(20):16, May 2011.
S. Flasche designed the experiment, analysed the data and wrote the manuscript. Presented in detail in section 3.1.1.

Stefan Flasche, Albert Jan van Hoek, Elizabeth Sheasby, Pauline Waight, Nick Andrews, Carmen Sheppard, Robert George, and Elizabeth Miller. Effect of Pneumococcal Conjugate Vaccination on Serotype-Specifc Carriage and Invasive Disease in England: A Cross-Sectional Study. *PLoS Med*, 8(4):e1001017, April 2011.
S. Flasche analysed the data and wrote the manuscript. Presented in detail in section 3.1.3.

Daniel M. Weinberger, Zitta B. Harboe, Stefan Flasche, J. Anthony Scott, and Marc Lipsitch. Prediction of Serotypes Causing Invasive Pneumococcal Disease in Unvaccinated and Vaccinated Populations. *Epidemiology*, 22(2):199207, March 2011.
S. Flasche contributed to the writing of the manuscript. Not presented in this Thesis

---

Stefan Flasche

July 30, 2012

Abstract

In September 2006 a 7-valent conjugate vaccine protective against pneumococcal carriage and disease was introduced into the national immunisation programme in the UK. This thesis assists with the assessment of the population impact of this vaccine and its 13-valent successor. The thesis links the observations in pneumococcal disease epidemiology to changes in carriage and assists with the understanding of the ecology of the pneumococcal serotypes.

Statistical models are employed to analyse the changes in post vaccination epidemiology in the UK in both carriage and invasive disease and to correct for possible ascertainment biases. Established mathematical models are analysed for their ability to reflect the pneumococcal ecology and are employed to describe the changes in post vaccination invasive disease. An individual based model is developed to account for the diversity of penumococcal serotypes and to study the likely mechanisms causing the distinct pneumococcal ecology, i.e. competition and coexistence.

Conjugate vaccination led to a decline in invasive disease associated with the targeted types and a consecutive increase of the non vaccine serotypes. Overall disease incidence declined while carriage rates stayed constant. This implies that the invasiveness of the replacing serotypes has been the main determinant for the impact of vaccination and that carriage data is essential to understand vaccine induced changes in disease.

Only one of the established models which describe competition amongst two pneumococcal serotype groups was found to commonly result in coexistence. This model could not fit the age patterns of Scottish post vaccination epidemiology, where serotype replacement amongst isolates of invasive disease was mostly confined to the elderly population. The individual-based model shows that serotype specific immunity can permit stable coexistence of a range of pneumococcal serotypes and that increasing the valency of a vaccine does not necessarily improve public health.

List of Abbreviations & Acronyms

CCR	Case:carrier ratio (invasiveness)
E&W	England and Wales
FOI	Force of infection
GAVI	The Global Alliance for Vaccines and Immunisation
HPA	Health Protection Agency
HPS	Health Protection Scotland
IBM	Individual-based model
ILI	Influenza like illness
IPD	Invasive pneumococcal disease
MLST	Multilocus sequence type
NVT	Non-vaccine types (ST not included in the vaccine)
ONS	Office for National Statistics
OR	Odds ratio
RR	Relative risk
ST	Serotype
VT	Vaccine types (ST included in the vaccine)
VE	Vaccine effectiveness

0 Thesis Outline

Streptococcus pneumoniae is a common bacterium which resides in the human nasopharynx and can cause severe disease including invasive pneumonia and meningitis. It is the single most important cause of childhood death in the developing world and is responsible for a high burden of disease in the developed world as well. Children are the main source of transmission of *Streptococcus pneumoniae* though the elderly population show high rates of invasive disease as well.

Vaccines conjugated to a carrier protein were found immunogenic in children and to be effective not only against progression to disease but also against acquisition of pneumoccal carriage. Thus, vaccination introduces indirect vaccine effects on the unvaccinated population (e.g. herd immunity). However, these conjugate vaccines offer protection only against a limited amount of the more than 90 different pneumococcal capsular serotypes. The vaccine targets the most prevalent serotypes amongst isolates of IPD. Hence, vaccination creates an ecological niche which is filled by the non-vaccine serotypes.

This thesis aims to assist with the analysis of the population impact of the introduction of the seven valent conjugate vaccine into the childhood immunisation scheme in Britain. The potential impact of the recently introduced PCV13 is also investigated. The mechanisms causing the observed complex ecological features pre and post vaccination are investigated. These mechanisms permit the circulation of the vast diversity of pneumococcal serotypes despite their competition for the same ecological niche. The role of immunity with these mechanisms is studied in detail.

The thesis is structured into seven chapters. In the first chapter common methods for the assessment of vaccine effects are presented. The value of dynamic models for estimates of the population impact of vaccination, which includes indirect vaccine effects, is discussed. Simple mathematical models, their extensions and

the threshold theorem for successful transmission of a pathogen are explained. Furthermore, the dependence of indirect effects of vaccination on the transmission potential of a pathogen is discussed alongside estimation techniques for such transmission potential.

The second chapter provides an overview of the pathogen *Streptococcus pneumoniae* and its structure. The role of its polysaccharide capsule for protection against the immune response is discussed, and evidence for the induced immune response following pneumococcal acquisition is reviewed. The chapter presents pneumococcal detection techniques and the surveillance systems for invasive pneumococcal disease in England & Wales and Scotland and discusses caveats for their interpretation.

Assessment of the population impact of 7-valent pneumococcal conjugate vaccination is dealt with in chapter three. The possible role of ascertainment in the increasing number of notifications for invasive pneumococcal disease in England and Wales prior to vaccine introduction is discussed and implications for impact analyses are drawn. Scottish data on invasive pneumococcal disease are analysed and compared to findings from England and Wales and other countries which have had the vaccine introduced to their national child immunisation schemes. Changes in pneumococcal carriage are studied by comparing data from a recent study, in children eligible for vaccination and their household members, to a similar study conducted before introduction of the vaccine.

In chapter four the aspects governing the transmission dynamics of the pneumococci are reviewed. Competition and coexistence of the serotypes are identified to be two main aspects of pneumococcal ecology which are poorly understood but likely to have a major effect on the indirect effects of vaccination. Hence, dynamic modelling approaches which include competition between two pathogens are first reviewed and then studied for their ability to provide meaningful and stable coexistence. Caveats of the model design as a result of artificially promoting coexistence are discussed.

In chapter five the most suitable of these dynamic models is then used to fit the observed population impact in Scotland following introduction of the 7-valent vaccine. Markov Chain Monte Carlo techniques are employed to fit the model to pre and post vaccination data and various data adjustments, which try to capture possible biases arising from changes in case ascertainment and temporal trends, were explored. This model was then transformed to an individual-based model

and expanded to assist the estimation of the impact of the 13-valent vaccine which was introduced in the UK in 2010.

In order to study the possible role of both heterologous and homologous immunity in governing competition and coexistence, and to develop a model framework free from the artificial promotion of coexistence, an individual-based model was developed in chapter six. A plausible parameter space, for pneumococcal carriage prevalence and the transmission potential of the serotypes is established. Implications for higher valency vaccination were drawn. These were further illustrated through studying the impact of competition on the threshold for successful vaccination in a simplified model.

The final chapter, seven, reviews the work in the thesis and discusses further research, how it would enhance current knowledge and how this research might be carried out.

1 Introduction

Outline

This chapter introduces some basic concepts about the evaluation of the effects of vaccination both before and after introduction of a vaccination programme into a countries' national immunisation programme. Several approaches for this are discussed. The value of mathematical modelling for an a priori assessment of indirect effects arising from a vaccine induced change in ecology is explained and an introduction to basic deterministic models for both epidemic and endemic pathogens is given. This includes the threshold theorem with the definition of a key parameter in modelling: the average number of secondary infections, the reproduction number. Its application for the estimation of a threshold where vaccination is sufficient to eradicate pathogen circulation is shown. Further, a few estimation techniques to derive the reproduction number directly from epidemiological data are explained.

1.1 The effects of vaccination

Vaccine effects are understood as the reduction in various disease outcomes associated with vaccination. Thereby one distinguishes between vaccine efficacy, effectiveness and impact. No universal definition is agreed on but in the following the most widely used nomenclature will be employed: Vaccine efficacy specifies the reduction in disease in vaccinated individuals measured by clinical trials (details follow); vaccine effectiveness characterises the same effect measured by various study designs in a partly vaccinated population and vaccine impact describes the effect of vaccination on a population level and therefore includes

Table 1.1: Nomenclature for the classification of individuals for the definition of the Risk Ratio

	Disease outcome	
	Yes	No
vaccinated	a	b
unvaccinated	c	d

factors such as vaccine coverage and indirect vaccine effects. A more detailed overview is provided by Halloran et al. [1]. The vaccine effect (efficacy, effectiveness or impact) ideally is measured by the relative risk (RR) which is the rate of disease outcome in one group (e.g. vaccinated individuals) compared to another group (e.g. unvaccinated individuals). Both groups should reflect a random sample from the targeted population. With the nomenclature given in table 1.1 the vaccine effect (Υ) is given by

$$\Upsilon = 1 - \mathrm{RR} = 1 - \frac{a/(a+b)}{c/(c+d)}. \tag{1.1.1}$$

However, for some study designs the relative risk cannot be computed. This is the case when study participants are selected on a basis other then exposure, in particular for subsequently discussed case control study designs participants are usually selected on the basis of disease outcome [2]. Therefore they do not represent a typical sample from the population and a relative risk cannot be computed (since the risk for disease amongst the groups does not reflect the risk in the population). In this case an odds ratio (OR) can be derived which provides a good approximation when the rare disease assumption holds (i.e. if the disease prevalence is low) [3]. Then, with the same nomenclature as before, the vaccine effect can be approximated through

$$\Upsilon \approx 1 - \mathrm{OR} = 1 - \frac{a/b}{c/d} = 1 - \frac{p_1(1-p_2)}{(1-p_1)p_2}, \tag{1.1.2}$$

where $p_1 = \frac{a}{a+b}$ and $p_2 = \frac{c}{c+d}$ are the probability for disease outcome amongst individuals from group 1 and 2 respectively. The odds ratio, although widely used, has a generally less intuitive interpretation than the relative risk and should therefore be treated with some caution.

One differentiates two major types of vaccines. Those which introduce immunity only against progression to disease but not against infection and those which also prevent acquisition of infection in the immunised individuals and therefore

introduce indirect effects by limiting the transmission of the disease; i.e. herd immunity (reduced infection rates in unvaccinated individuals through limited pathogen circulation in the population). Therefore, when estimating the effects of vaccination one needs to distinguish direct from indirect effects. The following provides a brief overview of some techniques to estimate the vaccine effects.

1.1.1 Estimation of direct effects

This subsection concerns techniques to measure vaccine induced direct protection of the vaccinated individual. This includes reduced disease outcome as well as a reduced propensity for asymptomatic infection.

Randomised controlled trials. Randomised controlled trials are considered to be the gold standard method for estimation of vaccine efficacy before the vaccine is licensed and widely deployed. In these, individuals are recruited randomly into two groups: one group will receive the considered vaccine and the other group will not (these will usually receive either a placebo or a vaccine targeting a different pathogen to avoid biases). These groups will be usually monitored for disease outcome (or sometime correlates for protection -e.g. antibody titres- are measured). This design allows for estimation of the vaccine efficacy by calculation of the Relative Risk as explained earlier.

Cohort study. There are generally two types of cohort studies: (i) prospective and (ii) retrospective designs. (i) Prospective cohort studies are relatively similar to randomised trials, i.e. they both follow up a cohort of vaccinated and unvaccinated individuals and the vaccine effectiveness can be inferred through the Relative Risk (i.e. the ratio of incidence in the vaccinated and unvaccinated cohort). However, in trials individuals are (randomly) assigned to receive the vaccine, while in a cohort study one defines the cohort members according to their vaccination status. Since the recruitment of individuals and their follow-up is extensive work and usually very costly other methods to estimate vaccine effects are pursued whenever possible. (ii) In retrospective cohort studies already collected data (e.g. from national surveillance when including personally identifiable information) is employed. The methods for estimation of vaccine effectiveness are the same as for prospective studies but no actual study is conducted, as it is rather a tool for posthoc analysis of existent data. For vaccines with

average coverage (which ensures that neither of the two groups is too small) and where disease presentation is rare (the study size would need to be very big) this is a beneficial option when data are available. For a more thorough discussion in particular of advantages and disadvantages of the approaches see e.g. Mann [4].

Case Control Study. When the disease outcome is rare, studies which select participants on the grounds of common exposure can be unfeasible. Here case control studies offer a way to approximate the incidence ratio of vaccinated and unvaccinated individuals by an odds ratio (compare equation 1.1.2). These are retrospective studies where besides individuals with the outcome of interest a disease free control group is chosen. The cases included should ideally be a random sample of all diseased individuals and while this can sometimes be challenging the main problem is the appropriate selection of controls. In the following three possible choices for controls are illustrated.

Test negative method. In the test-negative method, the control group is defined by those cases which were tested for an infection but found negative. This ensures a certain level of similarity of the control group and the group which was found test-positive, because both were symptomatic and presented to medical staff in order to get tested. However, a major problem for this approach is its dependence on the sensitivity of the test: with this choice of controls false-negative test results are considered controls which can introduce a bias if the test sensitivity is poor [5]. A recent example for the usage of the test negative method give Pebody et al. [6]; by employing data of the positivity of swabs taken by sentinel general practitioners from individuals presenting with influenza like illness during the pandemic and assuming 100% sensitivity, the authors found that the seasonal influenza vaccine did not decrease the risk of pandemic H1N1 2009 influenza.

Indirect cohort method. The indirect cohort method goes back to the work of Broome et al. in 1980 where the authors tried to estimate the effectiveness of a vaccine targeting a limited number of the pneumococcal serotypes [7]. This method has been slightly adapted to become a tool for case-control type analysis. It uses the cases presenting with disease caused by the same pathogen but subtypes not targeted by the vaccine as controls. This control group should represent a group of

individuals very similar to those presenting with vaccine-type disease. However, the method's applicability is limited to those few pathogens with vaccines available protecting against only a limited amount of subtypes. Further, if vaccination against some subtypes has effects on the remaining types (e.g. serotype replacement) this method can introduce a bias on the control group. Andrews et al. recently quantified this bias for the issue of serotype replacement following pneumococcal conjugate vaccination [8]. They found that the vaccine effectiveness assuming full replacement of vaccine targeted types with untargeted types is

$$\Upsilon = 1 - \left(1 + \frac{\Upsilon_c p_u}{1 - p_u}\right)(1 - \Upsilon_{icm}). \tag{1.1.3}$$

Here, Υ_c is the vaccine effect against acquisition of infection (carriage), $\Upsilon_{icm} = 1 - \text{OR}$ is the vaccine effectiveness as estimated for the unadjusted indirect cohort method, and p_u the proportion of infection due to vaccine types amongst unvaccinated individuals. Note, that if the vaccine protects exclusively against progression to disease given infection ($\Upsilon_c = 0$), no indirect effects are induced and the indirect cohort method is not biased by the choice of controls. Also, if the proportion of carriage of vaccine types is low amongst unvaccinated individuals the method introduces little bias. This might be the case as a result of strong herd immunity effects.

Screening method. In 1993 the screening method was introduced as a simple and rapid way to estimate vaccine effectiveness [9]. In this approach the whole population is used as a control group. It makes use of the symmetry of equation 1.1.2; i.e. the rate of disease in vaccinated and unvaccinated $p_{1,2}$ can be substituted with the vaccination coverage amongst the cases and the controls.

Post introduction incidence change. To measure the impact of vaccination the (disease, carriage, ...) incidence before vaccination is compared to the incidence after vaccination. Incidence here represents the risk for a defined disease outcome and in order to estimate the vaccine impact a relative risk can be computed similar to equation 1.1.1. The groups "vaccinated" and "unvaccinated" are hereby replaced by "individuals who have been eligible for vaccination in the post vaccination era" and "individuals in the pre

vaccination era who would have been eligible for vaccination in the post vaccination era". As eluded earlier the impact differs from the other assessments of vaccine effects in that it measures the population effect rather than the effect on individuals and therefore also includes factors like vaccine coverage and herd immunity effects; as such the impact is not a measure of only direct vaccine effects anymore. However, if high vaccine coverage is achieved vaccine impact, effectiveness and efficacy should be similar amongst vaccinated cohorts. Problems for this method include the variability of surveillance over time. Especially a relative risk based on a single years incidence estimate for each pre and post vaccination data might underestimate the temporal variation of both disease and surveillance; therefore pre vaccination data over a period of several years is essential for an appropriate analysis of the associated uncertainty.

1.1.2 Estimation of indirect effects

This subsection concerns techniques to measure vaccine induced indirect protection of the unvaccinated population. Vaccines which protect vaccinated individuals against infection can reduce the transmission of a pathogen in the population and thereby lead to reduced rates of infection in otherwise unprotected people; an effect usually referred to as herd immunity. However, other unforeseen indirect effects might arise from the interference with a pathogens ecology which poses the need for the consideration of indirect effects of vaccination when analysing a vaccine's benefit. In some cases indirect effects (positive or negative) might even outweigh the direct effects of vaccination.

Post introduction incidence change. As discussed above, for the direct effects of vaccination, estimating vaccine impact in the vaccine eligible population from post vaccination incidence change includes indirect effects as well. However, for estimating the indirect effects exclusively the changes in incidence in groups not eligible for vaccination need to be studied. This is done in a similar fashion to the assessment of the vaccine impact in the vaccine eligible population.

Cluster randomised trials Through a study design proposed by Halloran et al. direct, indirect and overall vaccine effects can be estimated [1, 10]. According to the anticipated vaccine uptake the study primary population (A) is

divided in two groups which either do or do not receive the vaccine. To be able to estimate indirect vaccine effects and the impact of vaccination from this trial another study population (B) needs to be defined. Ideally both populations should be separated in every way relevant to transmission dynamics to prevent indirect effects of vaccination in population A biasing population B. The indirect vaccine effect can then be estimated through the calculation of the relative risks for disease outcome amongst the unvaccinated individuals in population A and the individuals in population B. The vaccine impact can be determined through the relative risks of disease outcome in the whole of population A and population B.

Modelling. To estimate the potential population impact of vaccination before introduction of a vaccination programme into a country's immunisation schedule, and including both direct and indirect effects, the ecology of the respective pathogen has to be considered. This includes the likely impact of vaccination on the transmission dynamics in the vaccinated group and its effects on the unvaccinated individuals. Information including vaccine efficacy estimates, the likely uptake of the vaccination program and the targeted age- or risk-groups, the duration of vaccine protection, the population demographics, the routes of transmission and more generally the vaccine induced effect on the pathogens ecology need to be incorporated in order to sensibly estimate the population impact of a vaccination programme (see e.g. [11]). Mathematical modelling provides a flexible tool to do such analysis. In the following an introduction to simple modelling and the key aspects will be given.

1.2 Basic models and the role of the reproduction number

The theory of modelling for describing infectious disease transmission has a long standing history. It goes back to the 18th century when Daniel Bernoulli first defined parameters today known as the force of infection and the case fatality rate in his work on smallpox [12]. In 1927 the structure which most current epidemic models are still based on was published; the SIR or Kermack and McKendrick model [13]. This basic model defines a set of coupled non-linear differential equations which describe the epidemic spread of an infectious agent through

a randomly mixed population in dependence of the availability of susceptible individuals. Let S, I, R be the number of individuals being susceptible, infective and resistant for this agent, $N = S + I + R$ the population size, β the probability of an infectious contact between two individuals (the probability of contact of two randomly picked individuals times the probability for transmission given contact) and w the average rate of loss of infectiousness (this translates to $1/w$ being the duration of infectiousness under the assumption that it is exponentially distributed). Then the model equations are (time dependencies omitted):

$$\begin{aligned} S' &= -\beta IS \\ I' &= +\beta IS - wI \\ R' &= +wI. \end{aligned}$$

This is a very simplified representation of transmission dynamics but it represents the essential bit of theory, the dependence of new infections on the force of infection (FOI, $\lambda = \beta I$) and the available pool of susceptibles, in a mathematically accessible way. This approach takes several assumptions including that the duration of infectiousness is exponentially distributed, that the duration of infectiousness and infection is the same, homogeneity of the population (including that age differences have no role in the infection process and that each individual can meet another with the same chance), that the transmission dynamics are independent of external forcing and that infection induces lifelong immunity.

Assumptions on parameter distribution. While the assumption of generally exponentially distributed parameters in most cases does not represent reality there is little differences in model outcomes in most cases under different assumptions [14]. However, this assumption can be influential in some instances e.g. for antiviral treatment of infectious individuals during a influenza pandemic a short time after they became infectious and show symptoms [15]. The exponential distribution for time of infectiousness will cause the treatment to be efficient in reducing the time of infectiousness only for few individuals and therefore will introduce a potential bias towards the underestimation of the population effect of this treatment. This issue can be overcome e.g. by splitting the compartment I in half (see e.g. [16]). The duration of infectiousness will then be gamma distributed. If split further various Erlang distributions can be simulated (see e.g. supplement of [17]).

Infection and infectiousness. Differentiation between the compartments repre-

senting an infected and an infectious stage can be important in the case where the timing or duration of infection and infectiousness are different and one needs to fit to infection prevalence data. Assuming both to be similar would either alter the model transmission dynamics for which the duration of infectiousness is essential or bias the fitting process. Fortunately this differentiation can easily be achieved by adding compartment for periods where infectiousness and infection do not overlap. Similar things apply for the distinction of the symptomatic stage which is important for routine surveillance data.

Age differences. The immune system matures during the lifetime of an individual and introduces differences in the susceptibility of age groups to acquire an infection or in the time until the infection is cleared by the immune system. To incorporate this, the model can be split in multiple SIR models which represent the considered age groups and can have different durations of infectiousness. The age groups interact through the force of infection (FOI, λ); $\lambda = \sum_i \beta_i I_i$ incorporates the age dependent probability of an infectious contact and the number of infected persons for the respective age groups.

Random mixing. There are several approaches to avoid the assumption of random mixing of the population and reflect the heterogeneity of the population more adequately. Mixing according to age groups, as just discussed, is one of them which is often employed,as of data from contact surveys revealed that there are specific patterns of mixing which are usually focussed on increased mixing with similar age groups for industrialised countries [18]. Stratification of the population into different geographical areas can be done similarly to the age structuring. Data on daily commutes, the tracking of dollar bills or more abstract approaches have been used as proxy to reflect geographical mixing [19–21]. An additional area of research is the disease transmission through social networks which is probably the most appropriate reflection but also the one hardest to populate with data.

Immunity duration. In the SIR model by Kermack and McKendrick individials once infected with the disease become immune and don't lose their immunity. This leads to a steadily shrinking pool of susceptibles and eventually transmission ceases assuming no susceptible individuals are introduced by demographical changes (newborns or immigrants). For many pathogens

this is not the case; individuals might not establish immunity or are immune only for limited time. To reflect this SIS or SIRS models allow the waning of acquired immunity so that individuals become fully susceptible again.

External forcing. Many infectious pathogens follow a strong seasonal trend, such trends are not always well understood. For some pathogens seasonal variations are constrained to progression from carriage to disease (e.g. *Streptococcus pneumoniae*; see section 2.2) for others the seasonality also affects the transmission dynamics (e.g. influenza). While the loss of immunity could play a role in the repeated epidemics it is believed that some external seasonal forcing plays a major role (e.g. absolute humidity for influenza [22]). Mainly weather patterns have been associated with the seasonality of various pathogens, however few single factors have been found to significantly correlate with seasonal differences. Given this lack of understanding of the underlying mechanisms, seasonal differences are mostly modelled through sinusoidal waves which force the infection rate to comply with observed patterns (e.g. [23]).

Most models published to date employ a combination of these extensions to fit their purpose. However, adapted structures might be necessary i.e. when dealing with vector borne diseases or multiple pathogens which compete with each other.

From the simple SIR model important implications for the understanding and eventually the control of transmission dynamics can be derived. The arguably most important quantity for such models is the reproduction number R_0 (the average number of secondary infections a typical infected person causes during his period of infectiousness in a completely susceptible population). This is a dimensionless number which allows the comparison of the transmission potential of different pathogens and indicates the level of interventions needed for control. The threshold $R_0 > 1$ for a pathogen to spread effectively is self evident. This can be seen from the model equation in the change of the number of infectious individuals: if $I' > 0$ the number of infections in a totally susceptible population ($S = N = 1$) increases which corresponds to $R_0 \equiv \beta/w > 1$.

1.3 Threshold for successful vaccination

On the basis of the simple SIR model discussed in section 1.2 a simple threshold for successful vaccination can be derived. In order to stop the effective spread of the disease the effective reproduction number R_e (the average number of secondary infections a typical infected person causes during his period of infectiousness) needs to be below one. During an epidemic this is the case at its peak, caused by the substantial reduction in the number of susceptible individuals. The threshold for the reduction in the number of susceptibles achieved by vaccination to prevent the spread of the disease is: $R_e = \chi R_0 < 1$. Here χ represents the proportion of susceptible individuals amongst the population. Therefore if at least a proportion $1 - \chi = 1 - 1/R_0$ of the population is protected by either pre-existing immunity or vaccination (which itself is a function of vaccine coverage and vaccine effectiveness) no major outbreak will happen.

This theory also applies to models for endemic pathogens. The simplest model for such is an SIS model (note that an SIR model can be endemic as well, if an influx of susceptible people is permitted e.g. by introducing newborns to the population). With the same naming conventions as before the system of differential equations reads:

$$\begin{aligned} S' &= -\beta IS + wI \\ I' &= +\beta IS - wI. \end{aligned}$$

Here, w is the rate of loss of infectiousness ($1/w$ is the duration of infectiousness which is here assumed to be the same period as immunity to new infection). Endemic pathogens are considered to have a relatively constant level of infection amongst a population; in other words $R_e = 1$. In order to effectively vaccinate against such pathogens the same threshold as for epidemic pathogens $R_e < 1$ needs to hold. But in contrast to the epidemic model here individuals loose their infection induced immunity as they clear their infection. Therefore a proportion $1 - 1/R_0$ of the population needs to be protected solely by vaccination.

While in the SIS model described earlier failure to meet the vaccination threshold can simply lead to a failure of eradication but still a significant decrease in infections, this can have much wider implications. The vaccination threshold generally concerns only the indirect effects of vaccination; however these might outweigh the direct effects in some cases. If children were protected by a vaccine

which only provides protection for a limited number of years, and the uptake in combination with its effectiveness fails to meet the eradication threshold, a shift in age of infection can be the result. This poses a threat especially if this other age band is more likely to develop severe disease which is the case for rubella in pregnant women [24, 25].

The vaccine threshold for both SIR and SIS models in reality is a little more complex since e.g. vaccination of individuals with high contribution to the transmission dynamics has a greater population impact than vaccination of more isolated individuals. However, in essence the same threshold for the necessary reduction in transmission applies for more complex models.

1.4 Estimation of the reproduction number from epidemiological data

As discussed in the previous sections, R_0 is a key parameter for modelling a pathogen. Important information about its transmissibility and its threshold for vaccination can be derived from R_0. The definition of R_0 is based on a fully susceptible population, but in practise only the effective reproductive number $R_e = \chi R_0$ can be estimated (χ is the probability that a given contact is susceptible - this is equal to the proportion of susceptibles amongst the population in the simple SIR model). In some instance it can be reasonable to assume that all individuals are susceptible which means that $R_e = R_0$ but for the sake of precision the following will be about estimation of R_e. In order to infer R_e directly from epidemiological data several approaches exist. The methods can be split into those for epidemics/pandemics and endemics. A good overview is provided Heffernan et al. [26] which is referred to here unless mentioned otherwise.

1.4.1 Epidemics

Epidemics are characterised by an increasing rate of infections from usually low levels of prevalence which after a peak decline towards zero or previous levels again. Data from this early period of increase is often used to provide an early estimate for a pathogens transmissibility [27] but also other approaches exist. Some of the most widely employed methods are listed below.

Exponential growth rate. At the beginning of an epidemic the observed disease incidence will follow a (stochastic) exponential growth at rate r_0, i.e. $I' = r_0 I$, implying $I(t) = \mathrm{e}^{r_0 t}$. Given knowledge about the distribution of infectiousness during the infectious period of an individual one can infer R_e. In the simplest case, when infectivity is constant throughout the infectious period of length $1/w$ R_e can be estimated through

$$R_e = 1 + \frac{r_0}{w}. \tag{1.4.1}$$

Wallinga and Lipsitch [28] provide a good overview of how this estimate is derived and how different assumptions on the infectiveness affect the estimate. This includes the calculus for an empirical distribution. Let $a_0, a_1, \ldots, a_n$ be the the category bounds of the histogram defining the generation interval distribution and $y_0, y_1, \ldots, y_n$ the corresponding observed frequencies, then

$$R_e = \frac{r_0}{\sum_{i=0}^{n} y_i (\mathrm{e}^{-r_0 a_{i-1}} - \mathrm{e}^{-r_0 a_i})/(a_i - a_{i-1})}. \tag{1.4.2}$$

As an example, this proved helpful in the early stages of the influenza pandemic in 2009, when contact tracing provided data to infer an empirical distribution of the typical generation interval [27]. The major problem of this approach is the stochasticity of surveillance data and the likely changes in surveillance sensitivity during the early phase of an epidemic which makes the estimation of r_0 difficult.

Doubling time. A very similar approach concerns the estimation of the doubling time (D) of disease notifications to infer R_e. For the beginning of the epidemic where exponential growth can be assumed $r_0 = \frac{ln(2)}{D}$ holds (This can be easily seen from $\mathrm{e}^{r_0 t} = x$ and the need of r_0 to fulfill $\mathrm{e}^{r_0(t+D)} = 2x$). From the equation 1.4.1 when assuming exponential growth with a constant infectivity it follows that

$$R_e = 1 + \frac{ln(2)}{wD} \tag{1.4.3}$$

with the same constraints as before.

Final size. In contrast to the previous methods the final size method does not infer the reproduction number from early disease incidence data but gives an opportunity for post epidemic evaluation if the number of individuals who

escaped infection is known. Denote the number of susceptible individuals at the end of the epidemic by S_∞ then

$$R_0 = \frac{ln(S_\infty)}{S_\infty - 1}. \tag{1.4.4}$$

This assumes a SIR type model and the complete susceptibility of the population, however more general approaches can be derived. Due to the need for data on the number of infections, surveillance data is not sufficient and serosurveys need to be conducted to inform parameterisation using this approach.

Transmission tree inference. A very elegant method of inferring the effective reproduction number R_e from surveillance data employing a likelihood based approach was first described by Wallinga and Teunis [29]. A refinement of this method was later published by Cauchemez et al. and introduced a way how to reduce the bias of censoring towards the end of the considered period [30]. Hereby the relative likelihood p_{ij} that case i has been infected by case j, given their difference in time of symptom onset $t_i - t_j$ is expressed in terms of the probability distribution of the generation interval. Let $\phi(t)$ be the probability of a generation interval of length t. p_{ij} is then defined as the probability that case i has been infected by case j normalised by the probability that that case i has been infected by any other case:

$$p_{ij} = \frac{\phi\left(t_i - t_j\right)}{\sum_{k \neq i} \phi\left(t_i - t_k\right)} \tag{1.4.5}$$

and $p_{ii} \equiv 0$ (for the case that $\phi(0) \neq 0$). Now a case specific effective reproduction number can be derived by taking the sum over all cases weighted by the relative probability that case i has been infected by case j:

$$R_j = \sum_i p_{ij}. \tag{1.4.6}$$

From this the average effective reproduction number for any period of time can be calculated. A bootstrap approach provides a simple method to infer uncertainty around these estimates. For this a source case j for each case i is sampled according to the relative probabilities for infection p_{ij}. For each bootstrap simulation a set of $R_j \in mathbbN$ can be calculated by addition of the sampled number of secondary cases caused by case j. Amongst others this method was employed during the SARS outbreak and

the early phase of the 2009 H1N1 pandemic [27]. In an ideal world this method would identify the most likely transmission tree. In practice this would need complete data on infections and therefore is mostly applied differently: Disease notification data only represents a small fraction of the actual infections. If one assumes this fraction to stay similar over time the methods still gives sensible estimates. However direct links between individuals cannot necessarily be drawn. Extensions of the transmission tree inference method include the aggregation of some cases according to clusters (like school outbreaks or geographical compartments) or different groups of individuals [31].

1.4.2 Endemics

Endemics are characterised by a relatively stable number of infections over time. That is $R_e \approx 1$. Therefore, inference of R_e like for epidemic pathogens is not helpful and the assumption of a totally susceptible population which would yield $R_e \approx R_0$ is usually wrong in this case. However, a few methods exist to infer R_0:

Susceptibles at equilibrium. As shown in section 1.3 $R_e = \chi R_0$ holds with χ being the probability that a given contact is susceptible. Since $R_e \approx 1$ for endemics one can estimate R_0 through

$$R_0 = \frac{1}{\chi}. \tag{1.4.7}$$

This is a very simple method which has been used extensively. Assuming homogeneous mixing of the population, χ can be approximated through the proportion of susceptibles amongst the population (which is assumed to be at equilibrium for endemic pathogens). As for the other methods the precision of this approach hinges on the accuracy of the assumption of a homogeneous population and mixing patterns.

Average age of infection. This represents a closely related approach. Consider a pathogen (like varicella zoster) for which individuals are susceptible from birth and gain lifelong immunity once infected. Let A be the mean age at which an individual becomes infected and L the average life expectancy. If one assumes homogeneous mixing the average proportion of susceptibles in the model can be directly inferred by $\chi = \frac{A}{L}$. Hence, making use of the

equation 1.4.7 the reproduction number follows as

$$R_0 = \frac{L}{A}. \tag{1.4.8}$$

If the assumptions for this approach hold it offers a convenient way to estimate R_0 from readily available data.

Summary

In this chapter the basic methodology for the assessment of vaccine effects was discussed, including the use of relative risk and odds ratios and the terminology of vaccine efficacy, vaccine effectiveness and vaccine impact. Mathematical models have been identified to provide a useful tool to estimate indirect effects of vaccination before the countrywide introduction of a vaccine, where indirect effects are likely to arise as a result of reduced transmission. Simple modelling techniques to describe epidemic and endemic models are introduced and the key parameter which determines the transmission of a pathogen - the reproduction number- is defined. The threshold for successful vaccination is derived and several methods for estimating R_0 from surveillance or serology data are explored.

In the subsequent chapter an overview of background information on *Streptococcus pneumoniae* is given. Its impact on public health will be discussed along with its structure and resulting diversity. A summary of licensed vaccines targeting a variety of the pneumococcal serotypes and their induced protection on various endpoints (i.e. carriage and disease) will be provided. The chapter will further explain the routine surveillance systems in place to monitor pneumococci in England and Wales and Scotland along with the associated detection techniques and the caveats they introduce to any analysis. Not much is known about the influence of immunity induced by naturally acquired infection but evidence which is later used to evaluate the potential role of immunity in the ecology of the pneumococci is summarised.

Bibliography

[1] Halloran ME, Longini IM, Struchiner CJ (2009) Design and Analysis of Vaccine Studies (Statistics for Biology and Health). Springer, 1 edition.

[2] Deeks J (1998) When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses. British Medical Journal 317: 1155–6; author reply 1156–7.

[3] Greenland S, Thomas DC (1982) On the need for the rare disease assumption in case-control studies. American Journal of Epidemiology 116: 547–553.

[4] Mann CJ (2003) Observational research methods . Research design II : cohort, cross sectional, and case-control studies. Emergency Medicine Journal : 54–61.

[5] Orenstein EW, De Serres G, Haber MJ, Shay DK, Bridges CB, et al. (2007) Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. International Journal of Epidemiology 36: 623–31.

[6] Pebody R, Andrews N, Waight P, Malkani R, McCartney C, et al. (2011) No effect of 2008/09 seasonal influenza vaccination on the risk of pandemic H1N1 2009 influenza infection in England. Vaccine 29: 2613–8.

[7] Broome CV, Facklam RR, Fraser DW (1980) Pneumococcal disease after pneumococcal vaccination: an alternative method to estimate the efficacy of pneumococcal vaccine. The New England Journal of Medicine 303: 549–52.

[8] Andrews N, Waight P, Borrow R, Ladhani S, George R, et al. (2011) Using the Indirect Cohort Design to Estimate the Effectiveness of the Seven Valent Pneumococcal Conjugate Vaccine in England and Wales. PLoS ONE 6:

e28435.

[9] Farrington CP (1993) Estimation of vaccine effectiveness using the screening method. International Journal of Epidemiology 22: 742–6.

[10] Halloran M, Struchiner C, Longini IJ (1997) Study designs for evaluating different efficacy and effectiveness aspects of vaccines. American Journal of Epidemiology 146: 789–803.

[11] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[12] Dietz K, Heesterbeek JAP (2000) Bernoulli was ahead of modern epidemiology. Nature 408: 513–514.

[13] Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society 115A: 700–721.

[14] Conlan AJK, Rohani P, Lloyd AL, Keeling M, Grenfell BT (2010) Resolving the impact of waiting time distributions on the persistence of measles. Journal of the Royal Society, Interface / the Royal Society 7: 623–40.

[15] Lunelli A, Pugliese A (2008) Evaluating the effectiveness of antiviral treatment in models for influenza pandemic. Mathematical Medicine and Biology 25: 359–72.

[16] Dorigatti I, Cauchemez S, Pugliese A, Ferguson NM (2012) A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: Application to the Italian 20092010 A/H1N1 influenza pandemic. Epidemics 4: 9–21.

[17] Camacho A, Ballesteros S (2011) Explaining rapid reinfections in multiple-wave influenza outbreaks: Tristan da Cunha 1971 epidemic as a case study. Proceedins of the Royal Society 278: 3635–3643.

[18] Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine 5: e74.

[19] Hufnagel L, Brockmann D, Geisel T (2004) Forecast and control of epidemics

in a globalized world. Proceedings of the National Academy of Sciences 101: 15124–15129.

[20] Cooper BS, Pitman RJ, Edmunds WJ, Gay NJ (2006) Delaying the international spread of pandemic influenza. PLoS Medicine 3: e212.

[21] Danon L, House T, Keeling MJ (2009) The role of routine versus random movements on the spread of disease in Great Britain. Epidemics 1: 250–8.

[22] Shaman J, Goldstein E, Lipsitch M (2011) Absolute humidity and pandemic versus epidemic influenza. American Journal of Epidemiology 173: 127–35.

[23] Truscott J, Fraser C, Cauchemez S, Meeyai a, Hinsley W, et al. (2011) Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza. Journal of The Royal Society Interface .

[24] Knox EG (1980) Strategy for rubella vaccination. International Journal of Epidemiology 9: 13–23.

[25] Vynnycky E, White RG (2010) Infectious Disease Modelling. OUP Oxford, 1 edition, 137–141 pp.

[26] Heffernan JM, Smith RJ, Wahl LM (2005) Perspectives on the basic reproductive ratio. Journal of the Royal Society 2: 281–93.

[27] Ghani AC, Baguelin M, Griffin J, Flasche S, Pebody R, et al. (2009) The Early Transmission Dynamics of H1N1pdm Influenza in the United Kingdom. PLoS Currents Influenza : RRN1130.

[28] Wallinga J, Lipsitch M (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci 274: 599–604.

[29] Wallinga J, Teunis P (2004) Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol 160: 509–516.

[30] Cauchemez S, Boelle PY, Donnelly CA, Ferguson NM, Thomas G, et al. (2006) Real-time estimates in early detection of SARS. Emerging Infectious Diseases 12: 110–113.

[31] Glass K, Mercer GN, Nishiura H, McBryde ES, Becker NG (2011) Estimating

reproduction numbers for adults and children from case data. Journal of the Royal Society 8: 1248–59.

2 Background

Outline

The last chapter provided a brief introduction to the evaluation of vaccine effects. Various study designs to assess the direct effect of vaccination were discussed. Also the value of mathematical models in describing and predicting indirect effects like herd immunity was explained. The chapter further gave an outline of basic modelling techniques and assumptions, the threshold theorem including the definition of the basic reproduction ratio R_0 and possible refinements to relax various assumptions. A vaccination threshold in dependence of R_0 is derived for the basic SIR and SIS models which grants a sufficient level of indirect effects to cease the transmission in the whole population. Furthermore, various techniques to estimate the effective reproduction number from epidemiological data were reviewed.

This chapter provides background information on *Streptococcus pneumoniae* (*S. pneumoniae* or the pneumococcus) in order to better understand the epidemiological and ecological features that come with it. It provides the basis for the analysis presented in the following chapters. The structure of the pneumococcus and specifically the role of its polysaccharide capsule in evading the immune system and the progression to disease are discussed along with the variety of disease outcomes from pneumococcal infection. Further, current knowledge on naturally acquired and vaccine induced immunity against pneumococcal colonisation and progression to disease is reviewed and suitable correlates of immunity are explored. Currently there are three different types of vaccine approaches: the plain polysaccharide, the conjugated and the protein one. The vaccine formulations currently licensed or in development are described. Furthermore, detection techniques and routine surveillance in the UK are examined and the caveats (e.g.

the lack of detection of multiple carriage) they introduce to the data for the pneumococcus are studied.

2.1 The organism

Streptococcus pneumoniae is a gram positive bacteria which is the cause of a high burden of morbidity and mortality worldwide. In 2000, before respective vaccines were introduced to any national immunisation schedule, the pneumococcus was estimated to cause about 14.5 million episodes of serious pneumococcal disease in children younger than 5 years alone [1]. Furthermore, the pneumococcus was the likely cause of more than 800,000 annual deaths worldwide in children, aged 1-59 months, which corresponds to 11% of all deaths in this age band [1].

George Sternberg and Louis Pasteur discovered the existence of the pneumococcus simultaneously and independently in 1881 [2] but it took until 1974 until it was given its current name. The pneumococcus frequently colonises the human nasopharynx but the majority of these acquisitions do not result in disease. However, at times the penumococcus spreads locally or even invades the bloodstream or other sterile sites to cause severe disease.

The pneumococal outer cell surface is covered by a polysaccharide capsule which is believed to be the most important virulence factor since it protects the pneumococcus from phagocytosis [3]. Epidemiological evidence for the crucial role of the capsule for virulence was gathered in a study of invasive pneumococcal disease and nasopharyngeal carriage in children from Oxford [4]. Furthermore, a superior role of the capsule over the sequence type for association with mortality in Scotland was found by Inverarity et al. [5]. Reduced expression of the capsule gives greater accessibility to antibodies [6] and therefore eases clearance. However, the explicit associations between virulence and capsule are yet to be identified [7, 8]. There are more than 90 known different capsular serotypes (including types like 6C, 6D and 11E which have been discovered in the recent past) distributed over 46 serogroups [9–11]. However it is likely that the diversity of existing pneumococcal capsules is higher than currently anticipated. The cell wall is the layer underneath the pneumococcus' capsule. It consists of polysaccharides and teichoic acid and serves as an anchor for cell-wall-associated surface proteins [7]. The cell wall is believed to be protected by the pneumococcal polysaccharide capsule and causes the inflammatory reaction during pneumococcal infection [12].

Transmission of the pneumococcus is mostly assumed to be airborne via droplets [13–15]. However, it may also be passed on through contaminated surfaces e.g. hands. Once it has colonised the nasopharynx of the host the pneumococcus encounters the defensive mechanisms. Depending on both the ability of the hosts immune system to clear the invading pneumococcus and fight off infection and the pneumococcus' protection to clearance and its virulence *S. pneumoniae* either gets cleared which may take up to one and a half months [13] or progresses to disease.

2.2 Carriage and disease outcomes

The pneumococcus is a common resident of the human nasopharynx. In the absence of vaccination, carriage levels in children exceed 50% in most settings [16, 17] and are even higher in some native populations and developing world countries [18, 19]. The propensity to carry pneumococcus increases rapidly in the first few month of life [20] and then declines with age [21, 22]. Among other risk factors found to be associated (although not consistently among different studies) with pneumococcal carriage are the presence of siblings, day care center attendance, living in a rural area, exposure to passive smoke in the household and immunosuppression [21, 23]. Also prevalence is generally higher amongst indigenous populations and in developing countries [7, 24] which may result from an impaired immune system (see section 6.4). Season of the year was generally not found to influence the risk of pneumococcal colonisation [16, 25–27].

Following colonisation of the nasopharyngeal mucosal epithelium *S. pneumoniae* usually stays in the nasopharynx without progressing to disease until it is cleared by the immune system. However, pneumococcal carriage has the potential to result in various disease endpoints. At times the pneumococcus spreads locally and causes otitis media or sinusitis, spreads to the lungs and cause pneumonia or invades the blood stream to causes septicaemia (see figure 2.1) and in some instances meningitis. The progression from colonisation to disease is believed to occur within the first few weeks following acquisition of pneumococcal carriage [28, 29]. Therefore the event of acquisition of pneumococcal carriage rather then prevalence of carriage is important in relation to the progression to disease. However, *S. pneumoniae* is believed to transmit during the whole period of nasopharyngeal colonisation and therefore the duration of colonisation and carriage

prevalence is a relevant factor for its transmission.

A confirmed case of invasive pneumococcal disease (IPD) is defined as isolation of *S. pneumoniae* from a normally sterile body site (e.g., blood, cerebrospinal fluid, or, less commonly, joint, pleural or pericardial fluid) in a person of any age [30]. IPD is the most commonly and closely monitored disease endpoint of the pneumococcus and pools the severest outcomes of pneumococcal disease. As in carriage incidence (per population) of IPD is high in young children and declines from the age of one year onwards. However, despite continuously decreasing pneumococcal carriage prevalence in the elderly, rates of IPD increase and reach similar levels for 80+ years olds as observed for young children [31]. Despite little impact of seasonality on pneumococcal carriage, invasive pneumococcal disease in England and Wales follows strong seasonal patterns with a pronounced peak in late winter and few cases during summer [31]. However, seasonality differs in other climate settings [32–37]. Seasonal effects might be due to increased circulation of viral infections which leave the immune system prone to secondary bacterial infections. Correlation of the seasonal patterns of influenza and *S. pneumoniae* have been reported [36, 38], and influenza likely increases the susceptibility to progression from carriage to disease rather than enhancing general circulation of the pneumococcus (see section 2.7).

The pneumococcus is a major cause of morbidity and mortality worldwide. The highest child mortality rates are reported in central Africa and parts of Asia. Meningitis is the most severe disease endpoint of a pneumococcal infection with a global case-fatality rate of about 60% (about 35% in Europe, North America and Western Pacific and 65% in Southeast Asia and Africa). However, most deaths are attributable to pneumonia with death rates in Africa being more than 10-fold higher than in Europe [1]. The introduction of vaccines to national immunisation schedules across the developed world has further impaired this inequality leading to the development and distribution of a pneumococcal conjugate vaccine funded by the Global Alliance for Vaccines and Immunisation (GAVI) [39].

2.3 Pneumococcal vaccines

Before the era of vaccination against pneumococcus protection was provided solely by the human immune system and treatment relied on antibiotics. Nowadays four pneumococcal vaccines have been licensed for use by the European Commission

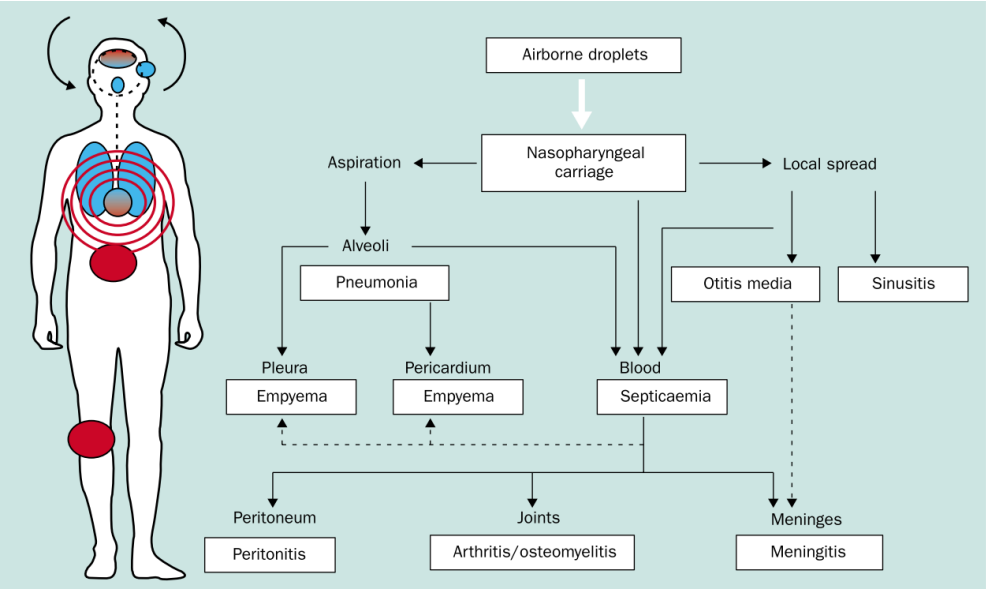


Figure 2.1: Pathogenic route for *S. pneumoniae* infection. Organs infected through the airborne and haematogenic routes are depicted in blue and red, respectively. From Bogaert et al [7].

and thus are available in the United Kingdom. PNEUMOVAX® 23 is generally recommend for use in the elderly population and in younger individuals at risk for pneumococcal disease and provides protection against disease caused by each of the 23 vaccine serotypes. By conjugation to a carrier protein the more recently licensed vaccines PCV7, PCV10 and PCV13 provide additional protection against colonisation with the vaccine types and are suitable for vaccinating children which are the main source for carriage and transmission of the pneumococcus. This induces partial indirect protection of the unvaccinated population through herd immunity but creates an ecological niche for non vaccine serotypes to exploit [40]. Hence, higher valency vaccines are being developed but the number of serotypes that can be included in a conjugated vaccine is limited [41]. Other approaches focus on more conserved antigens with a higher coverage amongst the pneumococci. Recently, the two highly conserved antigens PcsB and StkP were identified as suitable vaccine candidates since they were shown to play an important role in bacterial multiplication and could be found in 99% of all pneumococcal isolates collected from all over the world [42, 43]. A vaccine combining the conjugate approach with the protein candidate is currently ondergoing clinical trials [41].

2.3.1 Pneumococcal polysaccharid vaccine (PPV)

The pneumococcal polysaccharide vaccine PNEUMOVAX® 23 is manufactured by Merck. It was licensed in the United States of America in 1983 and designed to protect against disease caused by the capsular serotypes: 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F and 33F. The purified pneumococcal polysaccharides included in the vaccine induce antibody production which has shown to be effective in preventing pneumococcal disease [44]. Bacterial capsular polysaccharides induce antibodies primarily by T-cell-independent mechanisms. Therefore, antibody response to most pneumococcal capsular types is generally poor or inconsistent in children aged < 2 years - a group highly affected by pneumococcal carriage and disease - whose immune systems are immature [45]. Following the immunisation of pregnant women, antibodies are transferred to the neonate via the placenta as well as the breast milk [46–49]. However, clear evidence for actual protection of the newborns through this process has not been documented yet.

A meta-analysis from Melegaro and Edmunds [50] assessed the vaccine effectiveness (VE) of the 23-valent pneumococcal polysaccharide vaccine against pneumococcal pneumonia and invasive pneumococcal disease. For IPD in the general elderly population 2 studies were included in the analysis and for elderly at high risk 4 studies; for pneumonia 3 and 4 studies were included respectively. The authors found some evidence for PPV to offer protection against invasive disease in the general adult population aged 65 and older (VE = 65% [-49%, 92%]) and amongst those moderate effects in the high risk subgroup (VE_{risk} = 20% [-188%, 78%]). However, neither of the estimates was significantly different from zero. Protection against pneumococcal pneumonia was found to be generally weaker than protection against invasive disease. In the general elderly population the vaccine effectiveness was estimated to be 16% [-50%, 53%] and amongst those in a risk group the point estimate was negative (VE_{risk} = -20% [-92%, 25%]).

In 2003 the policy in the United Kingdom on PPV changed from a recommendation of vaccinating groups at risk (e.g. chronic heart disease, chronic lung disease, chronic liver disease, Diabetes mellitus, HIV, ...) to a general immunisation policy including all adults aged 65 years and over who have not previously been immunised. In England this change in policy was spread across three steps: From August 2003 onwards PPV was offered to all individuals aged 80 years and over, from April 2004 PPV was offered to all adults aged 75 and over and from April

2005 the program was extended to all adults aged 65 years and older; in Scotland from 2003 onwards PPV was offered to all individuals over 65 years of age.

2.3.2 Pneumococcal conjugate vaccines (PCV)

The pneumococcal conjugate vaccines are based on the conjugation of capsular polysaccharides to a carrier protein. The protein carriers then induce a T-cell dependent immune response. As with polysaccharide vaccines the conjugate vaccines induce protection only against the serotypes included in their formulation, although cross protection against a small group of serotypes, in particular 6A, has been reported (see section 3.1.3). In contrast to the polysaccharide vaccines the conjugate vaccines not only protect against invasive disease (or other disease outcomes) but also enhance serotype-specific clearance from the nasopharynx. Therefore conjugate vaccines reduce the circulation of the vaccine serotypes in the community and induce herd immunity. On the other hand they open an ecological niche and due to the ease of selective immunological pressure serotype replacement becomes an issue.

7-valent conjugate vaccine

The 7-valent pneumococcal conjugate vaccine (PCV7) Prevnar® is manufactured by Wyeth (now part of Pfizer). Its formulation covers the capsular serotypes 4, 6B, 9V, 14, 18C, 19F and 23F. Upon advice from the Advisory Committee on Immunization Practices (ACIP) the Centers for Disease Control and Prevention (CDC) recommended the vaccine administration to infants in the US in 2000 with a 3+1 schedule (recommended administration at 2,4 and 6 month and a booster dose at 12 month). The European Commission approved the marketing authorisation in February 2001 and the United Kingdom and many other European countries introduced PCV7 to their childhood immunisation schemes with some variation in schedule and uptake [51]. The UK introduced PCV7 into the national childhood immunisation programme from September 2006 in a 2+1 schedule with an additional catch up campaign for children up to two years of age.

10-valent conjugate vaccine

The 10-valent pneumococcal conjugate vaccine (PCV10) Synflorix® is manufactured by GlaxoSmithKline. In addition to the serotypes included in the 7-valent vaccine it covers serotypes 1, 5 and 7F. In January 2009 the European Medicines Agency (EMEA) recommended the granting of the marketing authorisation and in March 2009 Synflorix received authorisation by the European Commission. Using the non-typable *Haemophilus influenzae* (Hib) carriage protein D this vaccine is thought to provide additional protection against non-typable *Haemophilus influenzae* [52].

13-valent conjugate vaccine

The 13-valent pneumococcal conjugate vaccine (PCV13) Prevnar® 13 is manufactured by Wyeth (Pfizer) and is the successor of PCV7. In addition to PCV7 it covers the capsular serotypes 1, 3, 5, 6A, 7F and 19A. As in PCV7 all types are conjugated to the carrier protein CRM_{197}. In September 2009 the EMEA issued a positive opinion for PCV13 followed by the authorisation of the European Commission in December 2009. From March 2010 PCV13 has been administered in the US [53] and from April 2010 in the UK replacing the usage of PCV7.

2.4 Immunity

Great attention was drawn to improving the understanding of the complex mechanisms of immunity against *S. pneumoniae* in the last decade for the development of conjugate vaccination and also to assist with understanding the changes in ecology induced by vaccination. However, much uncertainty remains. Different aspects of immunity include the stratifications by innate vs. adaptive immunity, type specific vs. type unspecific immunity, immunity against pneumococcal carriage vs. immunity against disease and vaccine induced vs. naturally acquired immunity. The following will focus on the adaptive immunity stratified by vaccine induced and naturally acquired and will investigate evidence for both serotype specific and non serotype specific immunity.

2.4.1 Correlates of immunity

It is generally perceived that anticapsular Immunoglobulin G (IgG) antibody levels measured by an Enzyme Linked Immunosorbent Assay (ELISA) provide a robust correlate of existing protection (e.g. vaccine immunogenicity) against invasive pneumococcal disease [54]. A threshold of 0.35 $\mu g/ml$ is considered to offer full serotype specific protection [55, 56] although this might not be a robust measure for all capsular serotypes: evidence for ELISA being less sensitive for serotype 19F was found [57] and a recent study suggested that a threshold of about 0.2 $\mu g/ml$ might be more appropriate for 6B [58]. Many studies have suffered from the nonspecificity of the techniques used to measure titers of IgG levels in the early years [59] but by absorption with cell wall polysaccharide and 22F polysaccharide these issues could be overcome. These techniques are currently recommended in this form by the WHO as the gold standard [60–62]. The main protective mechanism against pneumococci is antibody mediated opsonophargocytosis, and therefore, not only the amount of antibodies but their functional activity was found to play an important role as a correlate of protection [63, 64]. Although, in general, in the investigated age groups good correlation was found between antibody levels and their functionality [57] this might not hold for the elderly population [65, 66].

The mechanisms for protection against carriage are less well studied. It was first suggested that increased levels of mucosal antibodies (IgA) are the most important corellate for protection against pneumococcal carriage since, as observed with *Haemophilus influenzae* type b, IgA levels were found to have increased after administration of the pneumococcal conjugate vaccine [67, 68]. However other studies suggested that colonization is prevented, not only by secretory pneumococcal anticapsular IgA, but by the presence, in saliva, of pneumococcal anticapsular IgG that either leaks from the serum or is secreted by committed plasma cells present on mucosal surfaces [69]. Following this hypothesis, Dagan et al. found serotype-specific IgG concentrations to negatively correlate with new acquisition events while no effect on the duration of carriage was detected [70]. Results from experiments in mice suggested that although being a correlate IgG concentration might not be the effector of protection but immunity against colonisation might be CD4+ T-cell dependent [71–73]. Evidence for this was found in humans as well [74]. Further studies suggested this pathway of protection accelerates the clearance rather then prevents acquisition [75].

As well as these biological markers, immunity against (type specific) pneumococcal carriage can be measured in the absence of antibody data. Longitudinal studies of pneumococcal carriage can provide further insights of the reduced likelihood of carriage given previous colonisation. As with the medical method this epidemiologic method suffers from caveats: the lack of detection of pneumococcal colonisation with multiple types (see chapter 2.5.2), and the possibility of the quickly maturing immune system in the early years of life biasing the interpretation of two swabs distant in time, are amongst them.

2.4.2 Vaccine induced immunity

The immunity induced by the polysaccharide vaccine (PPV) introduces an antigen response interacting directly with B cells inducing antibody synthesis in the absence of T-cells [76]. T-cell independent responses are restricted in numerous ways. Most importantly, they fail to induce significant and sustained amounts of antibody in young children below the age of 24 months which apart from the elderly represents the population at highest risk. While PPV is immunogenic in older children and adults the antibody response is dominated by IgM and IgG2. These are relatively short lived and hence do not support a booster response.

In contrast to plain polysaccharide vaccines the vaccines conjugated to a carrier protein induce an antibody response with an absolute requirement for T-cells. In consequence the response to protein antigens can be elicited in children under the age of 18 month. Also the immunity is long lived since immunological memory is generated. The antibody response is dominated by IgG1 and IgG3. The relative importance of memory versus circulating antibody levels for clinical protection by conjugate vaccines is unclear but it is interesting to note that even the least immunogenic of the Hib conjugates, PRP-D, has been shown to be efficacious in reducing the incidence of invasive Hib infection in Finland [77]. The efficacy of such formulations may thus be related to the ability of the conjugate vaccines to prime for memory, even with poor primary immunogenicity.

The protection induced by the vaccines is generally serotype specific by their design. However, because of the immunologic similarities it was anticipated that vaccine serotypes 6B and 19F will offer sufficient cross-protection against serotypes 6A and 19A respectively [78]. While there is good evidence that the response induced by the seven valent pneumococcal conjugate vaccine offers cross-

protection against carriage and disease caused by serotype 6A [79–81] protection against 19A is solely of theoretical nature so far [82, 83].

2.4.3 Naturally acquired immunity

Not much is understood about the mechanisms of natural acquired immunity against pneumococci. It is believed that pneumococcal colonisation induces an immune response which leads to development of anticapsular antibodies. Studies have found increased concentrations of IgA [84] and IgG [57] succeeding pneumococcal carriage.

While incidence for IPD generally follows a U-shaped age distribution with peak incidence occurring at similar levels in children and elderly [31] pneumococcal carriage is found to be high in young children and constantly declining by age thereafter [16, 22]. Therefore invasiveness (an inverse correlate for protection against disease given carriage) is thought to be high in youngster and declining in older children [85] and to increase again later in life.

In general there seem to be two major mechanisms governing natural immunity against the pneumococcus: strain-specific immunity and non strain specific immunity. These might play an important role for the ecology of the pneumococci and the potential contribution to it will be analysed in chapter 6. In the following evidence for both serotype specific immunity and serotype non specific immunity is presented.

Evidence for non strain specific immunity

Serotype independent immunity is generally believed to have two main sources: the general maturation of the immune system in the early years of life followed by its waning in elderly and immunity induced by type non-specific exposure.

As observed by Lipsitch et al. the serotype specific incidence of invasive pneumococcal disease, in the absence of vaccination, declines from about one year of age in a parallel manner. Given the vastly different circulation of these types, which results in different exposure, it seems evident that there needs to be an underlying mechanism of immunity which is not only strain independent but also increasing with age or rather accumulating with reinfection [86].

In a follow up study of 99 Bangladeshi children during their first year of life 435 identified episodes of pneumococcal carriage provided the basis for assessing the epidemiology of immunity in young children [20]. The authors employed a Cox proportional hazards model to estimate the effect of previous heterologous carriage on acquisition of the 4 most common serotypes and, adjusting for confounding factors, found that the probability for acquisition is significantly reduced by previous heterologous colonisation [87]. They could not detect any evidence that previous colonisation reduces homologous acquisition. However, the authors state that the critical age for maturation of the immune responsiveness to polysaccharide antigens is about 2 years and so the participants might not have been capable of producing anticapsular antibodies in response to pneumococcal carriage.

Challenging mice with live pneumococci of serotypes 6B, 7F or 14 was found to induce similar protection against recolonisation by homologous and heterologous serotypes [88]. Analogue tests performed in antibody deficient mice showed similar results suggesting that, in mice, protection against colonisation is antibody independent. They also reported the requirement of CD4+ T-cells for protection. Further animal studies suggested that antibody concentration correlates with protection against pneumococcal colonisation although not being functionally associated with it [71]. Also there is evidence that this mechanism of protection in mice reduces the duration of carriage rather then preventing acquisition [75]. However, how much of the findings in mice can be translated to humans is a matter of debate.

Evidence for strain specific immunity

At the beginning and the end of a 10 month longitudinal pneumococcal colonisation study conducted amongst 121 households in the UK [16] blood samples were taken from participants being 18 years or older. Capsular-specific IgG concentrations were measured using gold standard techniques. While the absence of carriage of a particular serotype in a household in the sampling interval lead to no or only a slight increase in the respective anticapsular IgG antibody concentration in an individual, detection of carriage in either the individual or in the household coincided with a significant increase in IgG levels for 4 out of the 6 analysed serotypes. Although not significant changes in IgG were estimated to be higher in previously exposed individuals than in those individuals naive for

serotype 18C. However, for serotype 6B no difference could be detected.

Between 2000 and 2004 more then 2,500 serum samples were collected in England and were analysed for the prevalence of age-specific, anticapsular pneumococcal antibody concentration (IgG) measured by ELISA [66]. The authors found great variety in IgG concentration by serotype and age. For six amongst eight analysed serotypes mean concentrations increased steadily with the participants age until the age of 20 years and stayed at a somewhat constant level thereafter. Profiles for seroypes 1 and 6b were different: while antibody concentration against serotype 1 was constantly high from 1 years of age, for 6B they increase till the age of 10-20 and decreased thereafter. These findings contradict the assumption of waning immunity with increasing age but could have been biased by offering PPV to the population over 65 years during this period and the uncertainty of IgG in being a sensible correlate of functional antibody avidity in elderly. Additionally, sample size in the elderly population was low that no fine stratification was possible and more pronounced age effects could have been hidden.

Epidemiological evidence for serotype specific immunity following pneumococcal carriage of the homologous type is sparse, primarily because of the lack of sufficiently large longitudinal studies. However, a longitudinal randomised trial was conducted in Israeli toddlers in day care, where subjects aged 12-35 month were followed up for 2 years. This provided a high prevalence setting for pneumococcal carriage with over 2000 samples being collected in a control group of unvaccinated individuals [70, 89]. Blood samples were obtained 0,12 and 24 month after enrolment but the determined serotype specific IgG concentrations suffered from the non-specificity of the applied ELISA. A generalized estimating equations model with logistic link was employed to estimate the association between prior carriage and risk of acquisition while adjusting for confounders like age and colonisation with other types [90]. The authors found evidence of serotype specific immunity following carriage of the homologous type for most of the considered serotypes although only in 3 out of 8 cases this was significant and the level of protection varied greatly. However no protection of prior heterologous carriage was detected.

2.5 Detection

2.5.1 Detection in disease isolates

Methods for detection of the pneumococcus from blood or other sterile sites is important for the clinical diagnoses and the appropriate treatment of patients with disease probably caused by *S. pneumoniae*. The available detection methods can be categorised into four major classes: 1) bacterial culturing techniques; 2)antigen detection; 3) demonstration of genetic material; and 4) serological antibody assays [91]. By far the most commonly used is bacterial culture. For this the fluid specimens are inoculated on 5% or 10% blood-agar and in 5% serum broth. The cultures are then incubated overnight and on the next day typical colonies of pneumococci may be observed. This may be confirmed by tests with optochin and bile. Details on standard culture media, reagents and the tests are provided by Lund et al. [92, 93]. A review on further development of culturing techniques is provided by the thesis of Kari S. Lankinen [91]. However, culturing pneumococcus suffers from poor sensitivity and only slow progress has been made in improving this [94].

2.5.2 Detection in carriage

Historically the method for detecting pneumococcal carriage was to inoculate a mouse with nasopharyngeal secretions and then culture the samples taken from the mouse for evidence of *S. pneumoniae* [95]. This was a labour intensive process which was unsuitable for testing a large amount of samples. Hence in the absence of a gold standard many other testing methods were employed and this variety introduced a possible bias into various studies [96]. In 1975, a detection method was published that showed no statistical difference relative to the isolation of pneumococcus through inoculation of mice (but found the mouse approach to be associated with higher sensitivity) [97]. *S. pneumoniae* was isolated by culturing nasopharyngeal secretions onto a blood agar plate containing gentamicin (similar to the method for recovering pneumococcus from sterile sites) which provided a cost saving and quick way of identification. However it took until 2003 until a gold standard for sampling and isolating pneumococcus was agreed on at WHO [24]. This recommendation is to sample pneumococcus from the nasopharynx using a calcium alginate or Dacron polyester-tipped swab. This sample should be transported in skim milk-tryptone-glucose-glycerin and cultured on blood (horse, sheep or goat) agar which contains 2.5 or $5\mu g$ gentamicin.

Although most pneumococcal carriage studies follow these recommendations and are relatively comparable, the study methods contain a few limitations. The procedure to sample from the nasopharynx is an unpleasant procedure for the sample donor and so it sometimes proves difficult to sample a sufficient fraction of the nasopharynx. Also it is difficult to detect pneumococcus when the nasopharynx is colonised in a low density which could potentially introduce a bias towards negative samples. More importantly, the pneumococcus is a diverse pathogen with over 90 different serotypes and the current gold standard method of growing cultures is limited in its capability of detection of co-colonisation in a sample. Recently developed methods show that this could lead to a significant underestimation of actual multiple colonisation. Turner and colleagues show that both the methodologies they tested (sweep method and detection through microarray) find considerably more serotypes in the same samples when compared to the WHO standard culturing method. While in 89% of 125 samples collected amongst Burmese/Karen refugee infants the culture method could only detect 1 serotype, but both other methods found multiple colonisation with 2-3 or occasionally even more serotypes in 43-48% of the isolates [98].

2.5.3 Typing

Idenfication of the serotype Differences in the degree of encapsulation are an essential determinant of the susceptibility to killing by neutrophils, and it is suggested that ultimately this defines the pneumoccus' success during nasopharyngeal carriage [8]. Furthermore, the capsule seems to be closely related to the invasive potential [4] and mortality [99] and, probably, carriage duration, resulting from the difference in susceptibility to killing [100]. The serotype of a pneumococcus is most commonly determined by quellung reaction which is the gold standard method. Quellung reaction or the Neufeld test is a biochemical reaction in which anticapsular antibodies bind to the capsule of the streptococcus together with India ink and cause the capsule to swell so that its contour becomes distinguishable with light microscopy [101]. In 2004 a latex agglutination test was developed, allowing accurate serotyping of some of the most common pneumococci without expertise in microscopy. Specifically, $10\mu l$ of latex suspension is mixed with $10\mu l$ of an overnight bacterial culture and a positive reaction is indicated by an agglutination appearing shortly after mixing. To minimize the number of agglutination tests the checkerboard system uses 14 latex pools [102].

However the quellung reaction has some limitations, including costs and labour, so both molecular and immunology-based serotyping methods have been pursued. Currently, the most promising assay combines an immunologic assay with multiplex PCR of serotype-specific genes [103]

Idenfication of the sequence type Although many characteristics of the pneumococcus were found to be mainly associated with its polysaccharide capsule [4,7,8] identifying genetic differences is valuable for monitoring events like capsular switching and the transfer of antibiotic resistant genes. Multi Locus Sequence Typing (MLST) characterises pneumococci by identifying sequence variation at seven housekeeping genes which encode essential metabolic functions.

2.6 Routine surveillance systems

There are numerous surveillance systems in place in the United Kingdom to monitor different aspects of pneumococcal disease. The most widely reported one is invasive pneumococcal disease. In the following the different surveillance systems for monitoring IPD in a) England and Wales and b) Scotland will be presented.

2.6.1 England and Wales

In England and Wales data on invasive pneumococcal disease is held in a reconciled dataset at the Health protection Agency (HPA). The reconciliation process involves two databases:

LabBase Weekly voluntary electronic reports of IPD are sent to the national database (LabBase) at the Health Protection Agency - Centre for Infections from hospitals and regional HPA units. About 5000-6000 reports per year of IPD are held in there. For the reconciliation process episodes are extracted if the flag for invasive was set to 'Yes' (i.e. the indicator for bacteraemia or meningitis was flagged or the Organism Patient Infection Episode (OPIE) record had the specimen type classified as sterile - i.e. blood). Only reports for England and Wales are included. Episodes of IPD are defined to be

distinct when they are more than 30 days apart from each other. De-duplication including the merging of non-distinct episodes is done.

MOLIS Hospitals and laboratories in England and Wales are asked to send their IPD isolates to the HPA Respiratory and Systemic Infection Laboratory (RSIL) for serotyping. This data is entered into a database called Modular Open Laboratory Information System (MOLIS). During the last decade an increasing amount of isolates were serotyped leading to approximately 4500-5000 reports entered into MOLIS in recent years. The same de-duplication proceedings as in LabBase are run for MOLIS as well.

Using the information on name (soundex), date of birth, hospital number, NHS number and specimen ID algorithms the MOLIS data are matched against the LabBase data to obtain a reconciled database.

2.6.2 Scotland

In 1999 an enhanced surveillance scheme to monitor invasive disease was introduced: the Scottish Pneumococcal Invasive Disease Enhanced Reporting (SPIDER). Every diagnostic laboratory in Scotland sends its reports on pneumococcal isolates to Health Protection Scotland (HPS) which collates the data jointly with the Scottish Meningococcus and Pneumococcus Reference Laboratory (SMPRL). The surveillance includes information on laboratory confirmation, serotype identification and antibiotic resistance profiling for blood and CSF isolates. Information on serogroup is available for all years whereas information on serotype only became available from 2003. Also since 2003 for all isolates the sequence type is determined and stored in SPIDER.

2.7 The pneumococcus and influenza

While the prevalence of pneumococcal colonisation is found to be rather stable throughout the season invasive pneumococcal disease follows strong seasonal trends with the main burden presenting during the winter months (see section 2.2). This implies that the transmission and circulation of the pneumococcus is not affected but the potential to progress to invasive disease is increased during winter. The reasons for this remain unknown. Possibly the immune system

of the host is weakened during that time. The mean weekly temperature and lack of ultraviolet radiation, and therefore reduced granulocyte and monocyte function and production of vitamin D in the host, was identified from a set of environmental variables (relative humidity, wind speed, atmospheric pressure and precipitation) as being associated with invasive pneumoccocal disease during a four year period in Philadelphia [34,104]. However, environmental factors, in particular absolute humidity, have been identified as relating to the seasonality of respiratory viruses, in particular influenza [105–108]. Respiratory infections (note that all present analysis is only done on reported incidence which introduces a strong bias towards the more severe respiratory illness episodes) could challenge the host's immune system in a way that provides the pneumococcus (and probably other pathogens) an opportunistic chance to invade. These associations were drawn in a number of studies [36, 109].

In the summer of 2009 pandemic A(H1N1) spread around the globe. First identified in Mexico, and estimated to pose a severe risk to public health [110], the strain spread to the United Kingdom [111]. Although transmission of the pandemic A(H1N1) in the UK was aseasonal, in contrast to many other European countries [108], and it was less severe than previously anticipated [112]. Similar aseasonality was reported in the United States and presented a unique opportunity to further investigate the relationship between influenza and the invasive pneumococcal disease [38]. The authors found strong spatial and temporal correlations of hospitalisation rates in ILI and IPD. Also the gradual decline in susceptibility by age which was observed for pandemic A(H1N1) was found to be mirrored by the relative increase in IPD.

In 2000, the Department of Health change the recommendation for routine vaccination of seasonal influenza in the UK [113]. While vaccination was still recommend for all individuals at increased risk of serious illness from influenza and for those in long stay residential accommodation, for the first time vaccination of people aged 65 years and over was advised and reimbursed by the government. Since then vaccination coverage rates have been relatively stable (see Baguelin et al. [114]). However, varying success in matching of the vaccine to the circulating strains might have influenced the effectiveness of the vaccine [115].

The Department of Health is currently reviewing routine vaccination against seasonal influenza. Studying seasonal influenza between the seasons 1995/96 and 2008/09 the impact of present vaccination programs was studied and the poten-

tial benefit of extension of this program was estimated through the use of dynamic models [114]. Findings suggest that routine vaccination of children could substantially reduce the burden of disease caused by influenza by reducing circulation in the age group believed to be the main reservoir for transmission and thereby also indirectly protecting other groups which are at high risk for severe disease. A cost-effectiveness study including different influenza associated disease outcomes confirms the public health benefit of this approach [116, 117].

Overall public health efforts to reduce the circulation of influenza haven't changed since the year 2000 and are therefore unlikely to have introduced a substantial effect on IPD incidence (and a bias in assessing the changes in the post vaccination era). However, if set in place, routine vaccination of children could considerably reduce circulation of influenza and therefore could complicate analyses of the impact of the 13-valent pneumococcal conjugate vaccine. These ecologic interdependencies need to be kept in mind.

Summary

S.pneumoniae is a major cause of worldwide morbidity and mortality amongst all age bands. The pneumococcus is the most common cause of preventable child-deaths which result from a variety of different diseases including meningitis and pneumonia. It is a highly diverse pathogen and the specific characteristics of its more than 90 types are mainly determined by their difference in capsule expression. Vaccines targeting a limited number of these capsules have been developed, licensed and introduced to an increasing number of countries. While the plain polysaccharide vaccine protects mainly elderly individuals from disease, with it's true efficacy being unclear, the vaccines conjugated to a carrier protein have proven immunogetic (using an IgG level of 0.35 $\mu g/ml$ as a threshold) in children, and efficacious in preventing colonisation, and therefore introducing the additional benefit of herd immunity. However, the true mechanism providing vaccine induced or naturally acquired immunity against pneumococcal colonisation remains unclear. From epidemiological studies there is evidence that following acquisition of pneumococcal carriage the host gains short-term protection against further acquisition from both homologous and heterologous types. This is an important feature for the ecology of the pneumococcus and has an impact on its epidemiology when following vaccination (see chapter 3). The work in chapter 6

will study the possible impact of immunity on the complex ecology and its implications. Detection of the pneumococcus in carriage and disease and the process of determining the corresponding serotypes has only recently been standardised. The currently used technique for serotyping is likely to considerably underestimate the number of types to be found from isolates. This provides a basis for the following analysis of the impact of the introduction of PCV7 into the national immunisation schedule in the United Kingdom (see chapter 3). The caveats introduced by the reporting systems and detection techniques will be discussed there. The inter-dependencies of influenza and IPD, while apparent, are unlikely to impact on the changes in pneumococcal incidence since 2000 but could introduce additional complexity if an extended vaccination schedule is implemented.

Bibliography

[1] O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, et al. (2009) Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. The Lancet 374: 893–902.

[2] Watson DA, Musher DM, Jacobson JW, Verhoef J (1993) A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. Clinical infectious diseases 17: 913–24.

[3] Watson DA, Musher DM (1999) A brief history of the pneumococcus in biomedical research. Seminars in Respiratory Infections 14: 198–208.

[4] Brueggemann AB, Griffiths DT, Meats E, Peto T, Crook DW, et al. (2003) Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. Journal of Infectious Diseases 187: 1424–1432.

[5] Inverarity D, Lamb K, Diggle M, Robertson C, Greenhalgh D, et al. (2011) Death or Survival from Invasive Pneumococcal Disease in Scotland: Associations with Serogroups and Multilocus Sequence Types. Journal of Medical Microbiology : 793–802.

[6] Magee AD, Yother J (2001) Requirement for capsule in colonization by *Streptococcus pneumoniae*. Infection and Immunity 69: 3755–3761.

[7] Bogaert D, Groot RD, Hermans PWM (2004) *Streptococcus pneumoniae* colonisation : the key to pneumococcal disease. The Lancet 4: 144–154.

[8] Weinberger DM, Trzccinski K, Lu YJ, Bogaert D, Brandes A, et al. (2009) Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathogens 5: e1000476.

[9] Henrichsen J (1995) Six newly recognized types of *Streptococcus pneumoniae*. Journal of Clinical Microbiology 33: 2759–2762.

[10] Park IH, Pritchard DG, Cartee R, Brandao A, Brandileone MCC, et al. (2007) Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. Journal of Clinical Microbiology 45: 1225–1233.

[11] Calix JJ, Nahm MH (2010) A new pneumococcal serotype, 11E, has a variably inactivated wcjE gene. Journal of Infectious Diseases 202: 29–38.

[12] Bruyn GA, van Furth R (1991) Pneumococcal polysaccharide vaccines: indications, efficacy and recommendations. European Journal of Clinical Microbiology & Infectious Diseases 10: 897–910.

[13] Melegaro A, Gay NJ, Medley GF (2004) Estimating the transmission parameters of pneumococcal carriage in households. Epidemiol Infect 132: 433–441.

[14] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, et al. (2010) Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC infectious diseases 10: 90.

[15] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[16] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[17] Huang SS, Hinrichsen VL, Stevenson AE, Rifas-Shiman SL, Kleinman K, et al. (2009) Continued impact of pneumococcal conjugate vaccine on carriage in young children. Pediatrics 124: e1–11.

[18] Hill PC, Cheung YB, Akisanya A, Sankareh K, Lahai G, et al. (2008) Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clinical Infectious Diseases 46: 807–14.

[19] Millar EV, O'Brien KL, Zell ER, Bronsdon MA, Reid R, et al. (2009) Na-

sopharyngeal carriage of *Streptococcus pneumoniae* in Navajo and White Mountain Apache children before the introduction of pneumococcal conjugate vaccine. The Pediatric Infectious Disease Journal 28: 711–6.

[20] Granat SM, Mia Z, Ollgren J, Herva E, Das M, et al. (2007) Longitudinal study on pneumococcal carriage during the first year of life in Bangladesh. The Pediatric Infectious Disease Journal 26: 319–324.

[21] Cardozo DM, Nascimento-Carvalho CM, Andrade ALSS, Silvany-Neto AM, Daltro CHC, et al. (2008) Prevalence and risk factors for nasopharyngeal carriage of *Streptococcus pneumoniae* among adolescents. Journal of Medical Microbiology 57: 185–9.

[22] Ridda I, Macintyre CR, Lindley R, McIntyre PB, Brown M, et al. (2010) Lack of pneumococcal carriage in the hospitalised elderly. Vaccine 28: 3902–4.

[23] Principi N, Marchisio P, Schito GC, Mannelli S (1999) Risk factors for carriage of respiratory pathogens in the nasopharynx of healthy children. Ascanius Project Collaborative Group. The Pediatric Infectious Disease Journal 18: 517–523.

[24] O'Brien KL, Nohynek H (2003) Report from a WHO Working Group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. The Pediatric infectious disease journal 22: e1–11.

[25] Gray BM, Turner ME, Dillon HC (1982) Epidemiologic studies of *Streptococcus pneumoniae* in infants. The effects of season and age on pneumococcal acquisition and carriage in the first 24 months of life. In: American Journal of Epidemiology. volume 116, pp. 692–703.

[26] Leino T, Auranen K, Jokinen J, Leinonen M, Tervonen P, et al. (2001) Pneumococcal carriage in children during their first two years: important role of family exposure. The Pediatric Infectious Disease Journal 20: 1022–1027.

[27] Syrjänen RK, Kilpi TM, Kaijalainen TH, Herva EE, Takala AK (2001) Nasopharyngeal carriage of *Streptococcus pneumoniae* in Finnish children younger than 2 years old. Journal of Infectious Diseases 184: 451–459.

[28] Gray BM, Converse GM, Dillon HC (1980) Epidemiologic studies of *Strep-*

tococcus pneumoniae in infants: acquisition, carriage, and infection during the first 24 months of life. Journal of Infectious Diseases 142: 923–33.

[29] Ghaffar F, Friedland IR, McCracken GH (1999) Dynamics of nasopharyngeal colonization by *Streptococcus pneumoniae*. The Pediatric infectious disease journal 18: 638–46.

[30] CDC (2010). Case definition for invasive pneumococcal disease. URL `http://www.cdc.gov/ncphi/disss/nndss/casedef/IPD_current.htm`.

[31] Trotter CL, Waight P, Andrews NJ, Slack M, Efstratiou A, et al. (2010) Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. Journal of Infection 60: 200–208.

[32] Syriopoulou V, Daikos GL, Soulis K, Michos A, Alexandrou H, et al. (2000) Epidemiology of invasive childhood pneumococcal infections in Greece. Acta Paediatra Supplement 89: 30–34.

[33] Roche PW, Krause V, Cook H, Barralet J, Coleman D, et al. (2008) Invasive pneumococcal disease in Australia, 2006. Communicable Diseases Intelligence 32: 18–30.

[34] White ANJ, Ng V, Spain CV, Johnson CC, Kinlin LM, et al. (2009) Let the sun shine in: effects of ultraviolet radiation on invasive pneumococcal disease risk in Philadelphia, Pennsylvania. BMC Infectious Diseases 9: 196.

[35] Ampofo K, Bender J, Sheng X, Korgenski K, Daly J, et al. (2008) Seasonal invasive pneumococcal disease in children: role of preceding respiratory viral infection. Pediatrics 122: 229–37.

[36] JANSEN AGSCGSC, SANDERS EAMAM, Van Der Ende A, VAN LOON AMM, Hoes AWW, et al. (2008) Invasive pneumococcal and meningococcal disease: association with influenza virus and respiratory syncytial virus activity? Epidemiology and Infection 136: 1448–1454.

[37] Brooks WA, Breiman RF, Goswami D, Hossain A, Alam K, et al. (2007) Invasive Pneumococcal Disease Burden and Implications for Vaccine Policy in Urban Bangladesh. American Journal of Tropical Medicine and Hygiene 77: 795–801.

[38] Weinberger DM, Simonsen L, Jordan R, Steiner C, Miller M, et al. (2011)

Impact of the 2009 Influenza Pandemic on Pneumococcal Pneumonia Hospitalizations in the United States. Journal of Infectious Diseases : 1–8.

[39] Moszynski P (2011) A pneumococcal vaccine is launched in Africa to cut child deaths. British Medical Journal 342: d1075.

[40] Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. (2003) Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. The New England journal of medicine 348: 1737–46.

[41] Rodgers GL, Klugman KP (2011) The future of pneumococcal disease prevention. Vaccine 29 Suppl 3: C43–8.

[42] Giefing C, Meinke AL, Hanner M, Henics T, Bui MD, et al. (2008) Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. The Journal of Experimental Medicine 205: 117–31.

[43] Nagy E (2010) IC47, a novel protein-based pneumococcal vaccine: from bench to the clinic. Tel Aviv, volume 4, p. 31. URL `http://www.ncbi.nlm.nih.gov/pubmed/7800387`.

[44] Austrian R (1977) Prevention of pneumococcal infection by immunization with capsular polysaccharides of *Streptococcus pneumoniae*: current status of polyvalent vaccines. Journal of Infectious Diseases 136 Suppl: S38—-S42.

[45] Centre for Disease Control and Prevention (1997) Prevention of pneumococcal disease: recommendations of the Advisory Committee on Immunization Practices (ACIP). Morbidity and Mortality Weekly Report 46: 1–24.

[46] Holmlund E, Nohynek H, Quiambao B, Ollgren J, Käyhty H (2011) Mother-infant vaccination with pneumococcal polysaccharide vaccine: persistence of maternal antibodies and responses of infants to vaccination. Vaccine 29: 4565–75.

[47] Almeida VDC, Mussi-Pinhata MM, De Souza CBS, Kubo CA, Martinez EZ, et al. (2009) Immunogenicity of 23-valent pneumococcal polysaccharide vaccine in HIV-infected pregnant women and kinetics of passively acquired antibodies in young infants. Vaccine 27: 3856–61.

[48] Quiambao BP, Nohynek HM, Käyhty H, Ollgren JP, Gozum LS, et al. (2007) Immunogenicity and reactogenicity of 23-valent pneumococcal polysaccharide vaccine among pregnant Filipino women and placental transfer of antibodies. Vaccine 25: 4470–7.

[49] Obaro SK, Deubzer HE, Newman VO, Adegbola RA, Greenwood BM, et al. (2004) Serotype-specific pneumococcal antibodies in breast milk of Gambian women immunized with a pneumococcal polysaccharide vaccine during pregnancy. The Pediatric Infectious Disease Journal 23: 1023–9.

[50] Melegaro A, Edmunds WJ (2004) The 23-valent pneumococcal polysaccharide vaccine. Part I. Efficacy of PPV in the elderly: a comparison of meta-analyses. European Journal of Epidemiol 19: 353–363.

[51] Gomes HDC, Muscat M, Monnet DL, Giesecke J, Lopalco PL (2009) Use of seven-valent pneumococcal conjugate vaccine (PCV7) in Europe, 2001-2007. Eurosurveillance 14: 1–6.

[52] Schuerman L, Borys D, Hoet B, Forsgren A, Prymula R (2009) Prevention of otitis media: now a reality? Vaccine 27: 5748–54.

[53] for Disease Control C, (CDC) P (2010) Licensure of a 13-valent pneumococcal conjugate vaccine (PCV13) and recommendations for use among children - Advisory Committee on Immunization Practices (ACIP), 2010. Morbidity and Mortality Weekly Report 59: 258–261.

[54] Siber GR, Thompson C, Reid GR, Almeido-Hill J, Zacher B, et al. (1992) Evaluation of bacterial polysaccharide immune globulin for the treatment or prevention of Haemophilus influenzae type b and pneumococcal disease. Journal of Infectious Diseases 165 Suppl: S129–33.

[55] Jódar L (2003) Serological criteria for evaluation and licensure of new pneumococcal conjugate vaccine formulations for use in infants. Vaccine 21: 3265–3272.

[56] Siber GR, Chang I, Baker S, Fernsten P, O'Brien KL, et al. (2007) Estimating the protective concentration of anti-pneumococcal capsular polysaccharide antibodies. Vaccine 25: 3816–26.

[57] Goldblatt D, Hussain M, Andrews N, Ashton L, Virta C, et al. (2005) Antibody responses to nasopharyngeal carriage of *Streptococcus pneumoniae*

in adults: a longitudinal household study. Journal of Infectious Diseases 192: 387–93.

[58] Goldblatt D, Southern J, Ashton L, Andrews N, Woodgate S, et al. (2010) Immunogenicity of a reduced schedule of pneumococcal conjugate vaccine in healthy infants and correlates of protection for serotype 6B in the United Kingdom. The Pediatric Infectious Disease Journal 29: 401–5.

[59] Soininen A, van Den Dobbelsteen G, Oomen L, Käyhty H (2000) Are the enzyme immunoassays for antibodies to pneumococcal capsular polysaccharides serotype specific? Clinical and Diagnostic Laboratory Immunology 7: 468–76.

[60] Concepcion NF, Frasch CE (2001) Pneumococcal type 22f polysaccharide absorption improves the specificity of a pneumococcal-polysaccharide enzyme-linked immunosorbent assay. Clinical and Diagnostic Laboratory Immunology 8: 266–72.

[61] Wernette CM, Frasch CE, Madore D, Carlone G, Goldblatt D, et al. (2003) Enzyme-Linked Immunosorbent Assay for Quantitation of Human Antibodies to Pneumococcal Polysaccharides. Clinical and Vaccine Immunology 10: 514–519.

[62] (2005). WHO global influenza preparedness plan. URL `http://www.who.int/csr/resources/publications/influenza/WHO_CDS_CSR_GIP_2005_5/en/`.

[63] Coughlin RT, White AC, Anderson CA, Carlone GM, Klein DL, et al. (1998) Characterization of pneumococcal specific antibodies in healthy unvaccinated adults. Vaccine 16: 1761–7.

[64] Johnson SE, Rubin L, Romero-Steiner S, Dykes JK, Pais LB, et al. (1999) Correlation of opsonophagocytosis and passive protection assays using human anticapsular antibodies in an infant mouse model of bacteremia for *Streptococcus pneumoniae*. Journal of Infectious Diseases 180: 133–40.

[65] Romero-Steiner S, Musher DM, Cetron MS, Pais LB, Groover JE, et al. (1999) Reduction in functional antibody activity against *Streptococcus pneumoniae* in vaccinated elderly individuals highly correlates with decreased IgG antibody avidity. Clinical Infectious Diseases 29: 281–8.

[66] Balmer P, Borrow R, Findlow J, Warrington R, Frankland S, et al. (2007) Age-stratified prevalences of pneumococcal-serotype-specific immunoglobulin G in England and their relationship to the serotype-specific incidence of invasive pneumococcal disease prior to the introduction of the pneumococcal 7-valent conjugate vaccine. Clinical Vaccine Immunology 14: 1442–1450.

[67] Bogaert D, Veenhoven RH, Ramdin R, Luijendijk IHT, Rijkers GT, et al. (2005) Pneumococcal conjugate vaccination does not induce a persisting mucosal IgA response in children with recurrent acute otitis media. Vaccine 23: 2607–13.

[68] Choo S, Zhang Q, Seymour L, Akhtar S, Finn A (2000) Primary and booster salivary antibody responses to a 7-valent pneumococcal conjugate vaccine in infants. The Journal of infectious diseases 182: 1260–3.

[69] Eskola J, Black S, Shinefield H (2003) Pneumococcal conjugate vaccines. In: Plotkin SA, Orenstein MD, eds. Vaccines. 4th ed. pp. 589–624.

[70] Dagan R, Givon-Lavi N, Fraser D, Lipsitch M, Siber GR, et al. (2005) Serum serotype-specific pneumococcal anticapsular immunoglobulin g concentrations after immunization with a 9-valent conjugate pneumococcal vaccine correlate with nasopharyngeal acquisition of pneumococcus. The Journal of infectious diseases 192: 367–76.

[71] Trzcinski K, Thompson C, Malley R, Lipsitch M (2005) Antibodies to conserved pneumococcal antigens correlate with, but are not required for, protection against pneumococcal colonization induced by prior exposure in a mouse model. Infection and immunity 73: 7043–6.

[72] Basset A, Thompson CM, Hollingshead SK, Briles DE, Ades EW, et al. (2007) Antibody-independent, CD4+ T-cell-dependent protection against pneumococcal colonization elicited by intranasal immunization with purified pneumococcal proteins. Infection and Immunity 75: 5460–4.

[73] Malley R (2010) Antibody and cell-mediated immunity to *Streptococcus pneumoniae*: implications for vaccine development. Journal of Molecular Medicine 88: 135–42.

[74] Zhang Q, Bagrade L, Bernatoniene J, Clarke E, Paton JC, et al. (2007) Low CD4 T cell immunity to pneumolysin is associated with nasopharyngeal

carriage of pneumococci in children. Journal of Infectious Diseases 195: 1194–202.

[75] Lu YJ, Gross J, Bogaert D, Finn A, Bagrade L, et al. (2008) Interleukin-17A mediates acquired immunity to pneumococcal colonization. PLoS Pathogens 4: e1000159.

[76] Goldblatt D (2000) Conjugate vaccines. Clinical and experimental immunology 119: 1–3.

[77] Peltola H, Kilpi T, Anttila M (1992) Rapid disappearance of Haemophilus influenzae type b meningitis after routine childhood immunisation with conjugate vaccines. Lancet 340: 592–4.

[78] Robbins JB, Austrian R, Lee CJ, Rastogi SC, Schiffman G, et al. (1983) Considerations for formulating the second-generation pneumococcal capsular polysaccharide vaccine with emphasis on the cross-reactive types within groups. Journal of Infectious Diseases 148: 1136–59.

[79] Whitney CG, Pilishvili T, Farley MM, Schaffner W, Craig AS, et al. (2006) Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. The Lancet 368: 1495–1502.

[80] Park SY, Moore MR, Bruden DL, Hyde TB, Reasonover AL, et al. (2008) Impact of conjugate vaccine on transmission of antimicrobial-resistant *Streptococcus pneumoniae* among Alaskan children. The Pediatric infectious disease journal 27: 335–40.

[81] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[82] Hausdorff WP, Hoet B, Schuerman L (2010) Do pneumococcal conjugate vaccines provide any cross-protection against serotype 19A? BMC pediatrics 10: 4.

[83] Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, et al. (2010) Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. Journal of Infectious Diseases 201: 32–41.

[84] Simell B, Kilpi TM, Käyhty H (2002) Pneumococcal carriage and otitis media induce salivary antibodies to pneumococcal capsular polysaccharides in children. The Journal of infectious diseases 186: 1106–14.

[85] Yildirim I, Hanage WP, Lipsitch M, Shea KM, Stevenson A, et al. (2010) Serotype specific invasive capacity and persistent reduction in invasive pneumococcal disease. Vaccine 29: 283–8.

[86] Lipsitch M, Whitney CG, Zell E, Kaijalainen T, Dagan R, et al. (2005) Are anticapsular antibodies the primary mechanism of protection against invasive pneumococcal disease? PLoS medicine 2: e15.

[87] Granat SM, Ollgren J, Herva E, Mia Z, Auranen K, et al. (2009) Epidemiological evidence for serotype-independent acquired immunity to pneumococcal carriage. Journal of Infectious Diseases 200: 99–106.

[88] Malley R, Trzcinski K, Srivastava A, Thompson CM, Anderson PW, et al. (2005) CD4+ T cells mediate antibody-independent acquired immunity to pneumococcal colonization. Proceedings of the National Academy of Sciences 102: 4848–4853.

[89] Dagan R, Givon-Lavi N, Zamir O, Sikuler-Cohen M, Guy L, et al. (2002) Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. Journal of Infectious Diseases 185: 927–936.

[90] Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, et al. (2008) Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. Journal of Infectious Diseases 197: 1511–8.

[91] Lankinen KS (2003) Catching the pneumococcus. Studies focusing on carriage, epidemiology and microbiological methods. Ph.D. thesis, National Public Health Institute / University of Oulu.

[92] Lund E (1960) Laboratory diagnosis of Pneumococcus infections. Bulletin of the World Health Organization 23: 5–13.

[93] Lund E, Henrichsen Jr, Hendrichsen J (1978) Laboratory diagnosis, serology and epidemiology of *Streptococcus pneumoniae*. In: Methods in microbiology, London: Academic Press, volume 12 of *Methods in Microbiology*. pp. 241–262. doi:10.1016/S0580-9517(08)70365-9. URL `http:`

//dx.doi.org/10.1016/S0580-9517(08)70365-9.

[94] Saha S, Darmstadt G, Naheed A, Arifeen S, Islam M, et al. (2010) Improving the Sensitivity of Blood Culture for *Streptococcus pneumoniae*. Journal of tropical pediatrics 0: 1–5.

[95] Austrian R (1986) Some aspects of the pneumococcal carrier state. The Journal of antimicrobial chemotherapy 18 Suppl A: 35–45.

[96] Tuomanen E, Mitchel T, Morrison D, Spratt B (2004) The Pneumococcus. Washington DC: American Society for Microbiology, 1 edition. URL http://books.google.com/books?id=ch6gBGP38dkC&printsec=frontcover&hl=de#v=onepage&q&f=false.

[97] Hendley JO, Sande MA, Stewart PM, Gwaltney JM (1975) Spread of *Streptococcus pneumoniae* in families. I. Carriage rates and distribution of types. The Journal of infectious diseases 132: 55–61.

[98] Turner P, Hinds J, Turner C, Jankhot A, Gould K, et al. (2011) Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. Journal of clinical microbiology .

[99] Lamb K (2010) Modelling Genetic Effects in the Transmission of Pneumococcal Carriage. Ph.D. thesis, University of Strathclyde.

[100] Melegaro A, Choi Y, Pebody R, Gay N (2007) Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. American journal of epidemiology 166: 228–35.

[101] Sorensen UB (1993) Typing of pneumococci by using 12 pooled antisera. Journal of clinical microbiology 31: 2097–100.

[102] Slotved HC, Kaltoft M, Skovsted IC, Kerrn MB, Espersen F (2004) Simple, rapid latex agglutination test for serotyping of pneumococci (Pneumotest-Latex). Journal of clinical microbiology 42: 2518–22.

[103] McEllistrem MC (2009) Genetic diversity of the pneumococcal capsule: implications for molecular-based serotyping. Future microbiology 4: 857–65.

[104] Dowell SF, Whitney CG, Wright C, Rose CE, Schuchat A (2003) Seasonal patterns of invasive pneumococcal disease. Emerging infectious diseases 9: 573–9.

[105] Shaman J, Kohn M (2009) Absolute humidity modulates influenza survival, transmission, and seasonality. Proceedings of the National Academy of Sciences 106: 3243–3248.

[106] Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M (2010) Absolute humidity and the seasonal onset of influenza in the continental United States. PLoS Biology 8: e1000316.

[107] Shaman J, Goldstein E, Lipsitch M (2011) Absolute humidity and pandemic versus epidemic influenza. American Journal of Epidemiology 173: 127–35.

[108] Flasche S, Hens N, Boëlle PY, Mossong J, van Ballegooijen WM, et al. (2011) Different transmission patterns in the early stages of the influenza A(H1N1)v pandemic: A comparative analysis of 12 European countries. Epidemics 3: 125–133.

[109] Kuster SP, Tuite AR, Kwong JC, McGeer A, Fisman DN (2011) Evaluation of Coseasonality of Influenza and Invasive Pneumococcal Disease: Results from Prospective Surveillance. PLoS medicine 8: e1001042.

[110] Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Kerkhove MDV, et al. (2009) Pandemic potential of a strain of influenza A (H1N1): early findings. Science 324: 1557–1561.

[111] Ghani AC, Baguelin M, Griffin J, Flasche S, Pebody R, et al. (2009) The Early Transmission Dynamics of H1N1pdm Influenza in the United Kingdom. PLoS Currents Influenza : RRN1130.

[112] Baguelin M, Hoek AJV, Jit M, Flasche S, White PJ, et al. (2010) Vaccination against pandemic influenza A/H1N1v in England: a real-time economic evaluation. Vaccine 28: 2370–84.

[113] Department of Health (2000). Major changes to the policy on influenza immunisation. URL `http://www.dh.gov.uk/en/Publicationsandstatistics/Lettersandcirculars/CMOupdate/DH_4003607`.

[114] Baguelin M, Flasche S, Camacho A, Demiris N, Edmunds WJ (2012) Reconstructing past influenza epidemics from consultation, virological surveillance data and a contact survey. (in preparation) .

[115] Jefferson T, Pietrantonj CD, Rivetti A, Bawazeer GA, Al-Ansary LA, et al. (2010) Vaccines for preventing influenza in healthy adults. Cochrane Database Systematic Review .

[116] Cromer D, Edmunds WJ, Jit M, Baguelin M, van Hoek AJ, et al. (2012) Estimating the burden of influenza by risk group. (in preparation) .

[117] Baguelin M, Flasche S, Edmunds WJ (2012) The cost-effectiveness of vaccination against seasonal influenza in the England. (in preparation) .

3 The impact of conjugate vaccination in Britain and elsewhere

Outline

In the previous chapter background information on *Streptococcus pneumoniae* was provided including an overview of the British surveillance systems in place and the methods for detection of IPD and pneumococcal carriage. The pneumococcus frequently colonises the nasopharynx, especially of children, and in some cases can progress to severe disease; mainly in children and the elderly. It is a very diverse pathogen with over 90 different serotypes distinguished by their polysaccharide capsule. A plain polysaccharide vaccine and conjugate vaccines with different valency do offer protection against some of the serotypes. While the polysaccharide vaccine is immunogenic only in adults, the conjugate vaccines stimulate the immune response in children as well, and do not only hinder disease progression but also prevent acquisition of infection with the serotypes included in the vaccine formulation. Through reduction in transmission of the vaccine types, indirect effects (changes in the epidemiology amongst those individuals not directly protected through vaccination) can be expected, including herd immunity effects, which is the protection of unvaccinated individuals through reduction in transmission, as an ideal result.

This chapter analyses the epidemiology of the pneumococcus in Britain and sets it in a global perspective. The first country to introduce PCV7 into the routine childhood immunisation scheme was the United States of America in October 2000 [1]. The program led to a substantial reduction of carriage prevalence of

the vaccine serotypes in the vaccinated age groups [2] and therefore to almost complete eradication of vaccine-type associated invasive disease [3]. Due to the herd immunity effect similar patterns were observed for age groups not eligible for vaccination. Little increase in non vaccine type (NVT) IPD (serotype replacement) was found following vaccination. Since the beneficial impact of introducing PCV7 was already observed shortly after introduction [1] many other countries subsequently introduced PCV7 to their childhood immunisation scheme; amongst them the United Kingdom where PCV7 was introduced in September 2006 in a 2+1 schedule (at 2,4 and 13 months). An additional one dose catch up program for children up to two years of age was set up to rapidly increase vaccine coverage. The effect of this vaccination campaign in Britain and the potential impact of higher valency vaccines, as well as a comparison of Britain with other countries' experiences are topics covered in this chapter.

Trends in invasive pneumococcal disease prior to the introduction of PCV7 into the national immunisation schemes of England & Wales and Scotland have been reported [4, 5]. To appropriately study the post vaccination epidemiology section 3.1.1 investigates whether a general increase in reporting could have induced the increase in the number of pneumococcal disease isolates prior to vaccination by comparing to control pathogens in England & Wales which similarly depend on blood culturing practice and have not been subject to any public health interventions. Assuming that the resulting conclusions can be transferred to potential changes in the Scottish IPD surveillance, the impact of PCV7 on invasive disease in Scotland is analysed in section 3.1.2 ,and a similar analysis for England & Wales is summarised. As outlined in section 2.2, nasopharyngeal colonisation precedes disease and only results in IPD occasionally. To detect the underlying changes in pneumococcal carriage, data collected amongst children and their family members in Hertfordshire is analysed in section 3.1.3. Conclusions about the likely impact of higher valency vaccines (PCV10 and PCV13) are drawn from this analysis. Finally the impact of PCV7 is set in a global context by reviewing the changes in epidemiology which resulted after it was introduced in different countries.

3.1 Pneumococcus in the UK

3.1.1 Pre vaccine trends

This research has been published [6].

Introduction To assess the impact of vaccination on the number of reported cases of IPD a method in widespread use is to compare post vaccination incidence directly with the pre vaccination levels (see section 1.1 and e.g. [3,7]). This implicitly assumes that disease incidence, in the absence of vaccination, had not changed, and that a similar level of ascertainment is maintained. Furthermore, is it assumed that there were no secular trends in individual serotypes (compare section 1.1: estimation of vaccine impact). A recent meeting hosted by the WHO, which reviewed the post-PCV7 data in different countries, identified changes in the sensitivity of the surveillance systems due to alterations in clinical awareness, reporting techniques and blood culturing practice as potential important confounders when making such comparisons over time and between countries [8]. The benefit of looking at invasive disease as an endpoint is that given its severity it is reasonable to assume that care seeking behaviour due to invasive pneumococcal disease is constant and at a high level. However identification rates and reporting might have increased due to increased awareness and advancing technology. Therefore rich pre vaccination data provides a useful source to determine trends which may bias estimates of the interventions impact.

Given the severity of IPD and the continuing universal access to the National Health Service (NHS) it is reasonable to assume that care seeking behaviour of IPD patients in England and Wales has remained constant in recent years. However, this might not be the case for laboratory investigation or reporting behaviour which may have been subject to changes in practice over time. Reporting rates for IPD in hospitalised cases have been shown to vary with blood culturing rates which may have changed as clinical practice has evolved [9], while recent technical developments may have improved reporting of laboratory-confirmed cases [10].

To appropriately interpret changes in the incidence of IPD after introduction of PCV7 in England and Wales and to infer likely similar implications for Scotland the potential role of changing reporting sensitivity is assessed by evaluating the following hypothesis.

Hypothesis If there was a trend of increasing numbers of notifications of invasive pneumococcus disease prior to the vaccine introduction date (VID), an analogous trend should be found in notifications of control pathogens which depend similarly on blood culturing practice and reporting and which have not been subject to any interventions. Furthermore, if this was true, by studying the course of incidence in the control pathogens in the years after VID one could infer the likely role of reporting biases to the changes in epidemiology of IPD in the same period.

Methods The control pathogens selected for comparison with IPD to test this hypothesis were those of the most commonly reported bacteraemias that fulfil the following criteria: (i) they have to be endemic in England & Wales and are not solely outbreak related, (ii) they should be mainly diagnosed through blood culture, (iii) they should not have been subject to vaccination or any other public health intervention during the period of comparison, (iv) and they should provide statistically robust numbers in each of the considered age bands. The pathogens identified using these criteria were *Escherichia coli* and non-pyogenic streptococci (*Streptococcus acidominimus*, *S. bovis*, *S. gordonii*, *S. intermedius*, *S. mitis*, *S. mutans*, *S. oralis*, *S. parasanguinis*, *S. salivarius*, and not further typable: alpha-haemolytic *Streptococcus sp.*, non-haemolytic Streptococcus sp., *S. anginosus* group, *S. milleri* group, *S. mitis* group, *S. sanguinis* group).

Data on disease episodes of these pathogens reported between July 2000 and June 2010 (until June 2007 for *S. pneumoniae* - and respectively stratified in epidemiological years) and identified by blood culture were obtained from the national routine laboratory surveillance system of England and Wales (LabBase - see section 2.6.1). To account only for whole disease episodes additional isolates with similar OPIE (see section 2.6.1) identifier were excluded from the analysis when reported less than 14 days apart. As part of the enhanced surveillance episodes of IPD were checked for duplicates using personal identifiers. The periods pre and post VID where approximated by the epidemiological years until 2006/07 and from 2007/08. Hence, the post VID period includes two month where PCV7 hasn't yet been introduced.

The data in all age groups for all pathogens was over dispersed relative to a Poisson distribution (dispersion parameter > 2.5) hence for each of the pathogens a negative binomial regression model was fitted to the observed number of cases (λ_t) per 100,000 population (modelled by a population offset n_t). The model includes a linear trend and a five level factor (γ_j) to indicate the pre VID years

($i = 1$) and the post VID years: 2006/07, 2007/08, 2008/09, 2009/10 ($i = 2 \dots 5$).

$$ln\left(\frac{\lambda_t}{n_t}\right) = \alpha + \beta t + \sum_{j=2}^{5} \gamma_j I(t = t_j)$$

Here $t \in T = \{t_1, \dots, t_5\}$ is indicating the pre and post VID years as defined above and $I(\cdot)$ is the indicator function which is 0 if the expression in the brackets is false and 1 otherwise. The (anti-logged) slope β of this model indicates the pre VID trend; i.e. the annual percentage change in the rate $e^{\beta} > 1$ indicates a positive trend and a slope $e^{\beta} < 1$ a negative trend. Each of the post vaccination factors indicates the age adjusted deviation (relative risk) of the post VID data from the extrapolated pre 2005/06 trend (compare impact of vaccination in section 1.1).

To test for differences in slopes over pathogens a corresponding model for the pre-vaccination era was employed:

$$ln\left(\frac{\lambda_t^*}{n_t}\right) = \alpha + \beta t + \sum_{j=2}^{3} \psi_j I(P = P_j) + \sum_{j=2}^{3} \phi_j t : I(P = P_j).$$

Here λ_t^* holds the observed number of cases for *Escherichia coli*, IPD and non pyogenic streptococci, and $P \in \{P_1, P_2, P_3\}$ represents the different pathogens. The ψ_j estimate the differences in prevalence between the three pathogens and the ϕ_j are included to determine the significance of the interaction between time and pathogen (i.e. to test for difference in the increase in incidence over time prior to pneumococcal conjugate vaccination between the pathogens).

For this analysis individuals were stratified into three age groups (<5, 5-64, 65+). Information on age was missing in 1623/176660 (1%) of all *Escherichia coli*, in 452/34830 (1%) of all non pyogenic streptococci and 441/43646 (1%) of all IPD cases. Due to the small numbers this was not modelled as an offset, but the number of cases in each age group was inflated proportionally to the pathogen specific age distribution in each year:

$$\lambda_{t,a} = \lambda_{t,a}^1 + \lambda_t^2 * \frac{\lambda_{t,a}^1}{\sum_a \lambda_{t,a}^1}$$

Here a represents the age group and λ_t^1, λ_t^2 the number of cases at time t which have or have not age information, respectively.

All analysis was performed in R Version 2.11.

Table 3.1: Number of cases reported for *S. pneumoniae* and the control pathogens, by age group

		00/01	01/02	02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10
S. pneumoniae	<5	470	465	554	540	518	558				
	5-64	1352	1235	1629	1759	2051	2155				
	65+	2161	1924	2310	2301	2274	2138				
Escherichia coli	<5	269	296	336	414	361	427	408	458	468	470
	5-64	3159	3198	3718	4457	4532	4956	5152	5755	5898	6188
	65+	8280	8395	9517	11236	10964	12143	12140	13241	13871	15205
non pyogenic	<5	208	233	286	331	293	349	329	413	437	428
streptococci	5-64	785	875	946	1288	1187	1301	1334	1389	1395	1356
	65+	812	880	908	1194	1183	1293	1277	1463	1432	1445

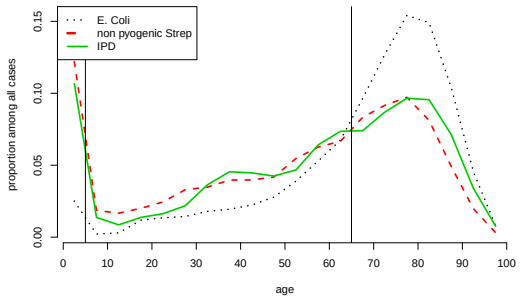


Figure 3.1: Age distribution of the three pathogens over all ages between 2000/01 and 2005/06, inclusive.

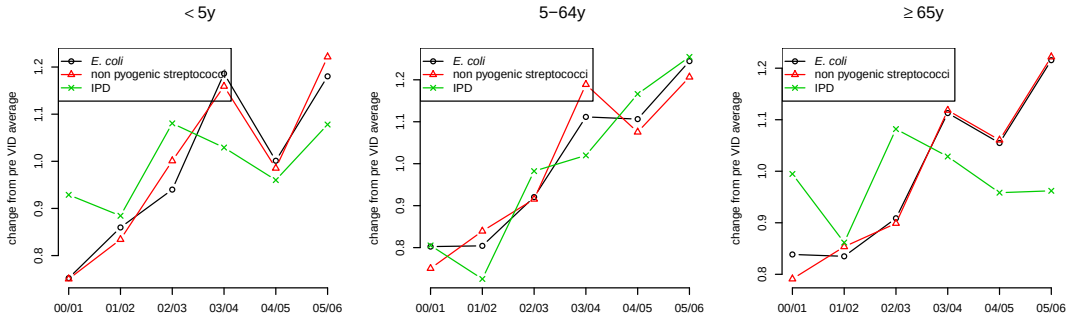


Figure 3.2: Pre 2005/06 trends in invasive pneumococcal disease and control infections stratified by age group. The incidence is presented standardised to the pathogen specific pre 2005/06 average number of cases for comparability.

Results In total more then 150,000 disease episodes were considered in the pre VID era. Those were differently distributed amongst the age bands (see figure 3.1). While IPD and non-pyogenic streptococci showed similar patterns, most of the *Escherichia coli* episodes were reported in the elderly population. In the population under five years of age 27% of disease episodes were due to *Escherichia coli*, 49% to IPD and 25% to non-pyogenic streptococci. In the 5 to 64 year-olds the respective distribution was 50%, 34% and 16%, and for those 65 years and older it was 70%, 21% and 9%.

In the age bands under 5 years and between 5 and 64 years IPD incidence showed a positive trend. This trend was most pronounced in the 5 to 64 age group with an estimated yearly increase of about 11% ($p<0.001$). In the under five year olds reports of IPD incidence increased 3% ($p= 0.031$) per year. In over 65 year old no significant trend ($p=0.91$) was detected (see table 3.2).

For *Escherichia coli* and non pyogenic streptococci significantly positive trends were estimated for all age bands (all $p<0.001$). In the 5 to 64 year olds pre VID

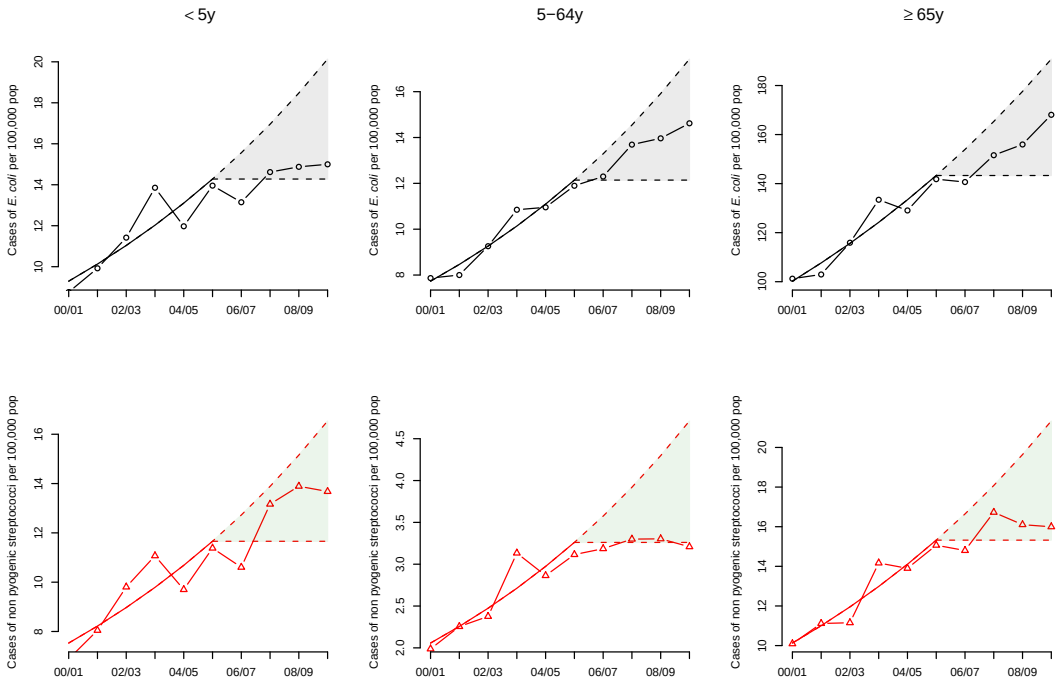


Figure 3.3: Pre- and post-VID trends in control infectiouns stratified by pathogen and age group. The solid line represents the respective pre-vaccination model fit and the grey shaded era the prediction interval which is bounded by the extrapolated pre vaccination trend and the predicted incidence in 2005/06 (upper and lower dotted lines respectively).

Table 3.2: Estimates for the relative rates of the pre vaccination trend models

	IPD	*Escherichia coli*	Non pyo Strep
<5	1.030 [1.002,1.058]	1.091 [1.052,1.132]	1.093[1.051,1.138]
5-64	1.106 [1.075,1.138]	1.095 [1.075,1.114]	1.097 [1.061,1.134]
65+	0.999 [0.972,1.027]	1.074 [1.055,1.193]	1.087 [1.063,1.111]

trends in IPD were not significantly different from these in *Escherichia coli* and non pyogenic Streptococcus. Significant differences between pre VID trends in IPD compared to both *Escherichia coli* and non pyogenic Streptococcus were found in the population under five years of age ($p<0.05$) and 65 years and older ($p<0.001$). However differences in the point estimates in the children were smaller than in the elderly (see table 3.2).

In the youngest age group the reported incidence of non-pyogenic streptococci closely followed the prediction extrapolated from the trend before the season 2006/07 (figure 3.3). However, the prediction for *Escherichia coli* significantly overestimated the actual incidence (see table 3.3 although the reported incidences were still higher then before the date of PCV7 introduction. In the 5 to 64 year olds the pre-PCV7 trend of non-pyogenic streptococci diminished in the seasons after introduction of PCV7 and the incidence remained at the level of the incidence observed in the season 2005/06. Again *Escherichia coli* differed: the observed incidences were in between the predictions that assumed the pre-PCV7 trend to continue and those assuming no trend after introduction of PCV7. In the oldest age group the reports for both *Escherichia coli* and non-pyogenic streptococci were in between these two prediction scenarios.

Discussion The findings suggest that the pre VID trends in the age group 5 to 64 years in invasive pneumococcal disease are similar to those observed in *Escherichia coli* and non pyogenic Streptococcus and that these trends continued in part. While the extrapolated trends were not entirely consistent with the incidence observed in the control pathogens after introduction of PCV7 this may nevertheless provide a likely indication of the expected incidence of bacteraemia reports post-PCV7 compared to those assuming the trend to discontinue after 2005/06.

The reason why there is no positive pre vaccination trend in the 65+ population in IPD but there is a trend in the other pathogens is unknown. One reason could be different reporting of these pathogens in elderly. Another possibility

Table 3.3: Percentage difference between the incidence of control infections and their incidence predicted by continuing the pre-vaccination trend (Model A) and discontinuing it (Model B)

		Escherichia coli		non pyogenic Strep	
		Model A	Model B	Model A	Model B
<5	06/07	-15.5% [-27.2,-2.0]	-8.1%	-16.7% [-27.2,-2.0]	-9.2%
	07/08	-13.8% [-26.8,1.6]	5.1%	-5.2% [-19.6,11.8]	15.9%
	08/09	-19.5% [-33,-3.2]	8.7%	-8.3% [-23.9,10.4]	24.3%
	09/10	-25.5% [-39.4,-8.4]	9.5%	-17.3% [-32.9,2.0]	22.3%
5-64	06/07	-7.4% [-14.1,-0.2]	4.4%	-10.9% [-22.5to 2.6]	0.7%
	07/08	-5.9% [-13.4,2.4]	15.9%	-15.8% [-28.0,-1.5]	4.0%
	08/09	-12.3% [-20.1,-3.6]	18.0%	-23.1% [-35.6,-8.3]	3.9%
	09/10	-16.1% [-24.4,-6.8]	23.3%	-31.9% [-44.1,-17.1]	-0.8%
65+	06/07	-8.6% [-15.4,-1.3]	2.3%	-11.1% [-18.7,-2.8]	0.7%
	07/08	-8.3% [-15.9,0.0]	10.8%	-7.5% [-16.2,2.1]	14.3%
	08/09	-12.2% [-20.3,-3.3]	15.0%	-18.0% [-26.7,-8.3]	11.0%
	09/10	-11.9% [-20.9,-1.9]	25.2%	-25.0% [-33.8,-15.0]	11.5%

could be the effect of the pneumococcal polysaccharide vaccine (PPV) which was introduced to the population aged 65 years and more subsequently from August 2003.

The employed data sources introduced some limitations to the analysis. While for *Escherichia coli* and non-pyogenic streptococci data from the national laboratory-based surveillance were employed, the data on *S. pneumoniae* was further enhanced and checked for duplicates. This was necessary due to inconsistencies specific to IPD in the national laboratory-based surveillance dataset in the years 2002 and 2003 when duplicates had been included in error. Such duplicates were removed from the enhanced IPD dataset (which contains an extra $10-20\%$ of cases identified solely from referral of isolates for serotyping). This might introduce an ascertainment bias when comparing these notification data from slightly different sources. A further limitation which may have affected the data was the migration in mid-2001 of the national surveillance database to a new platform which could have caused inconsistencies for all pathogens if data deduplication was compromised. Non-pyogenic streptococci are sometimes associated with contamination. This could artificially increase the reported incidence of bacteraemias caused by non-pyogenic streptococci. However, this is unlikely to alter the trend estimates for increasing ascertainment since this would be equally reflected by the contaminated samples.

Secular trends in *S. pneumoniae* in the absence of vaccination were reported [11,12]. These trends are poorly understood, cannot be predicted and, if occurring

during the period under investigation, are likely to affect the estimates of the vaccine impact; i.e. the extrapolation of the overall trends in the pre VID period assumes not only a continuing trend in ascertainment but continuing secular trends as well. Although this assumption is likely only to be true in the short term we show that in the younger age groups the other pathogens had similar positive pre VID trends which suggest that trends caused by common source were more pronounced in that period than any secular trends since these trends are unlikely to be similar for the different pathogens.

The positive trends pre VID of pneumococcus and the other pathogens in the age groups <5 and 5-64 could reveal a common source causing this trend, which is likely to be a matter of increased ascertainment rather than a parallel change in disease prevalence. This could be due to numerous reasons including increasing blood culturing practice and an increasing number of laboratories choosing to report these non-notifiable diseases to the national database [13]. Additional factors might have contributed to the observed trends, such as increasing automation of detection techniques or improved survival of people with underlying conditions, which could have increased the numbers of vulnerable people in the population.

To assess the probable development of reported cases of IPD in the absence of vaccination two predictions, continuing pre-PCV7 trend (Model A) and no trend (Model B) were compared to the actual reports for *Escherichia coli* and non-pyogenic streptococci. Although no clear evidence was found that one of the prediction models was superior, the reported number of cases was in between both predictions in all age groups. While Model A might provide the better prediction for the under five year-olds, Model B seemed to provide more reliable estimates in the age group between five and 64 years.

These findings have important implication for analysing the effect of the introduction of PCV7 to the childhood immunisation scheme: By ignoring the pre vaccination era trend one would underestimate the reduction in IPD and overestimate serotype replacement. However, accounting for these trends one risks overestimating the effect of PCV7 and underestimating the impact of serotype replacement. This analysis helps to estimate the uncertainty introduced by changing ascertainment when analysing the effects of PCV in England and Wales. Similar analyses from other countries would be helpful to improve the comparability of the vaccine effects. In the absence of such data the following analysis of IPD in

Scotland assumes similar implications arising from the pre vaccination increase in IPD incidence.

3.1.2 Scotland

This research is in preparation for submission as part of a general assessment of pneumococcal conjugate vaccination in the UK.

Introduction In September 2006 the 7-valent pneumococcal conjugate vaccine was introduced into the Scottish Routine Childhood Immunisation Programme as a 2+1 schedule for infants and a single dose catch-up campaign for children up to two years of age. Overall 100,226 children in Scotland were involved in the catch up programme. By the December 31st in 2007 86.1% of all children had completed the PCV7 vaccination appropriately for their age [14]. All cases of IPD in Scotland have been routinely monitored by Health Protection Scotland (see section 2.6.2). Data on specimens tested positive for pneumococcus by blood culture or from cerebrospinal fluid from week 27 in 2000 to Week 26 in 2010 were analysed for post-vaccination changes in vaccine type (VT) and non VT (NVT) disease, as well as in specific serotypes.

This section evaluates the impact of introducing PCV7 to the national childhood immunisation scheme in Scotland. It addresses potential caveats which might mask the true impact. Secular trends of some serotypes, including serotypes 1 and 5, have been reported previously [12]. The proportion of serotype 1 among all IPD in Scotland was found to have significantly increased in the pre-PCV era [5,15]. Hence, including serotype 1 in the analysis might lead to biases in the evaluation. Hence, the results will be presented both with and without inclusion of serotype 1. A possible role for increased reporting of IPD incidence in England and Wales was found (section 3.1.1) which might introduce a bias in the post vaccination epidemiology in Scotland in a similar fashion. Thus two approaches, one adjusting for pre vaccination trends and one not, were applied.

Methods Data from the Scottish national surveillance from 2000/01 to 2009/10 on invasive pneumococcal disease isolates obtained from blood or cerebrospinal fluid were extracted. 27 (0.4%) cases were omitted from the analysis due to missing information on their age. For 637 cases (10.1%) no information on the

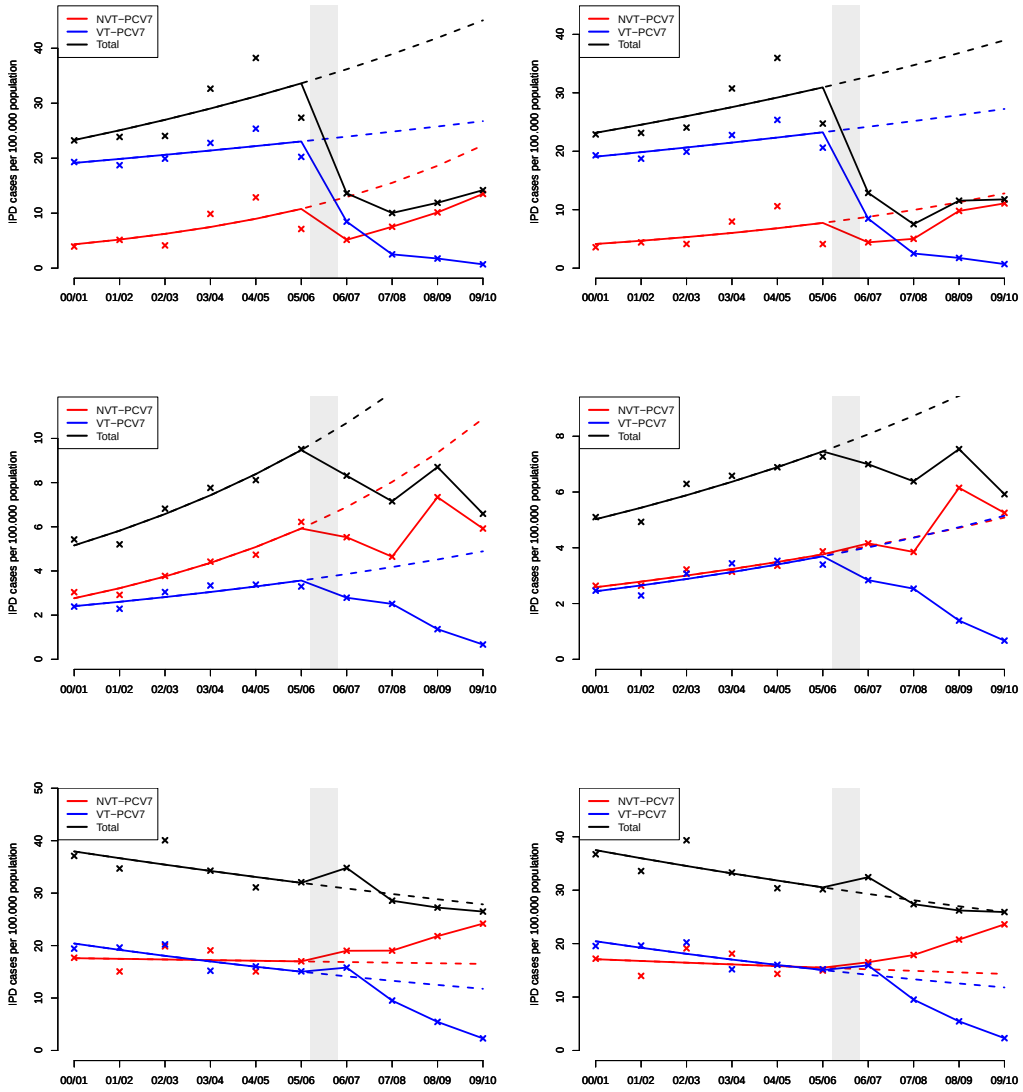


Figure 3.4: IPD incidence in VT and NVT disease in children less than 5 years of age, 5 to 64 year olds and over 64 year olds (top to bottom) including Serotype 1 (left column) and not including Serotype 1 (right column). The crosses represent the data on invasive pneumococcal disease, the solid lines represent the fitted model and the dashed lines show the extrapolated pre vaccination trend.

serotype was available. The numbers for the analysis where therefore inflated proportionally according to the year- and age-specific serotype distribution observed (as done in section 3.1.1). Further, pre 2003 for most of the isolates only information on serogroup was available. Therefore, for those isolates with only information on the serogroup the serotype distribution within a serogroup was assumed to be similar to the one observed in the respective age group between 2003 and the introduction of the vaccine and were distributed accordingly. The impact of the vaccination program was estimated in two different ways: (method 1) by comparing the average incidence in the year 2004/05 and 2005/06 to the incidence in 2009/10, parametric bootstrap techniques were applied to estimate 95% confidence intervals and (method 2) by predicting the post-vaccination incidence in the absence of vaccination allowing for a trend in the pre-vaccination years. For the latter a Poisson regression model was fitted to the observed number of cases (λ_t) a similar to that employed in section 3.1.1: an intercept α, a linear trend (β), a population offset (n_t) and a five level factor (γ_j) to represent the pre-vaccination period and the post-vaccination years 2006/07, 2007/08, 2008/09, 2009/10.

$$ln\left(\frac{\lambda_t}{n_t}\right) = \alpha + \beta t + \sum_{j=2}^{5} \gamma_j I(t = t_j)$$

A Poisson model was chosen because the data was not found to be highly over dispersed. The dispersion parameter was estimated to be 2.6 at most (for NVT in younger age groups) and was smaller when excluding serotype 1 from the analysis. The analysis was stratified into three age groups: under 5 years of age, 5-64 years old and over 64 years old. Serotype specific analysis using the same model was done only for serotypes accounting for at least 1% of IPD incidence: 1, 3, 4, 6A, 6B, 7F, 8, 9N, 9V, 11A, 12F, 14, 18C, 19A, 19F, 20, 22F, 23F, 33F. This analysis was corrected for multiple testing by the Bonferroni correction [16] (n=7 for PCV7 types and n=12 for non-PCV7 types) and not stratified by age groups due to small numbers.

To test whether the pre-vaccination slopes of VT ($i = 1$) and NVT ($i = 2$) were similar a model including and interaction term similar to the one in chapter 3.1.1 was employed:

$$ln\left(\frac{\lambda_t}{n_t}\right) = \alpha + \beta t + \psi_2 I(P = P_2) + \phi_2 t : I(P = P_2)$$

where $P_j \in \{VT, NVT\}$.

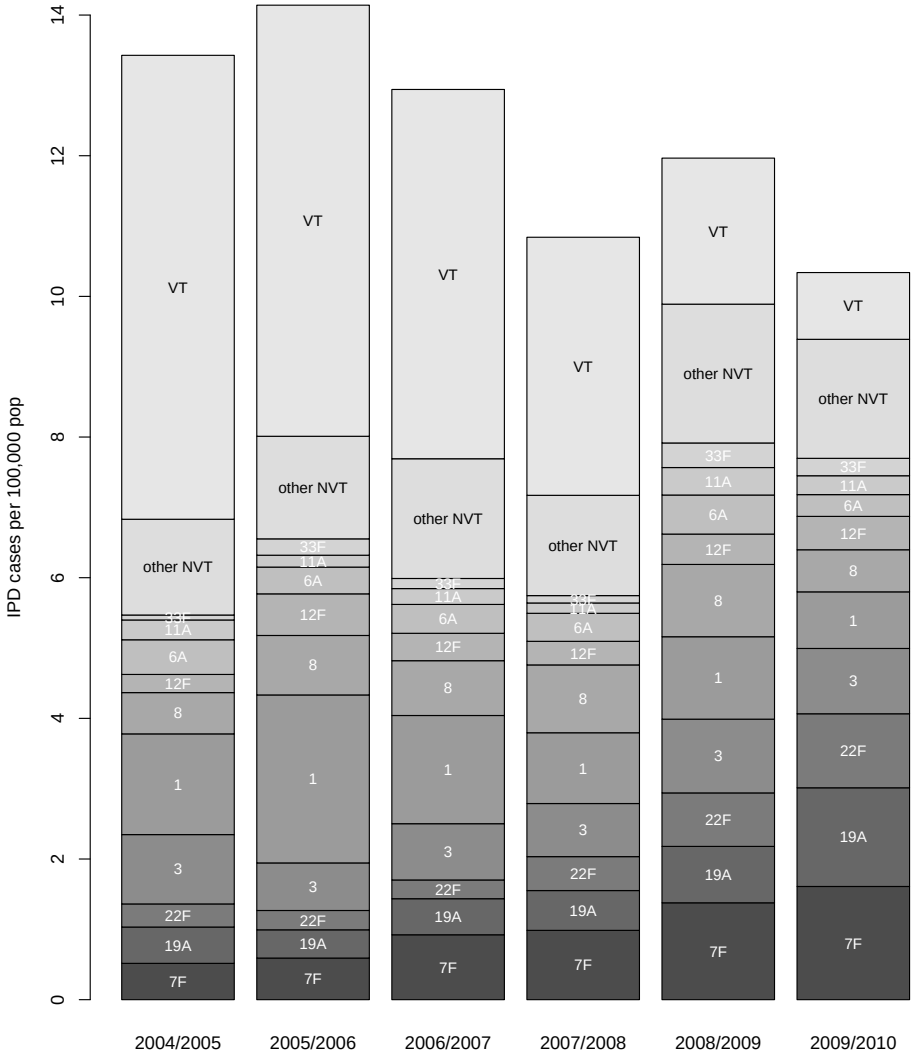


Figure 3.5: Changes in non-vaccine serotype IPD incidence distribution for all age groups in the epidemiological years 2004/05 to 2009/10

Results 7219 episodes of invasive pneumococcal disease were included in the analysis. The overall IPD incidence increased steadily until 2005/06 and gradually decreased thereafter.

After the introduction of PCV7, by 2009/10, the incidence of IPD caused by serotypes included in the vaccine had declined by 97.0% with method 1 and 97.4% with method 2 in children less than 5 years of age (table 3.4). In the age groups 5

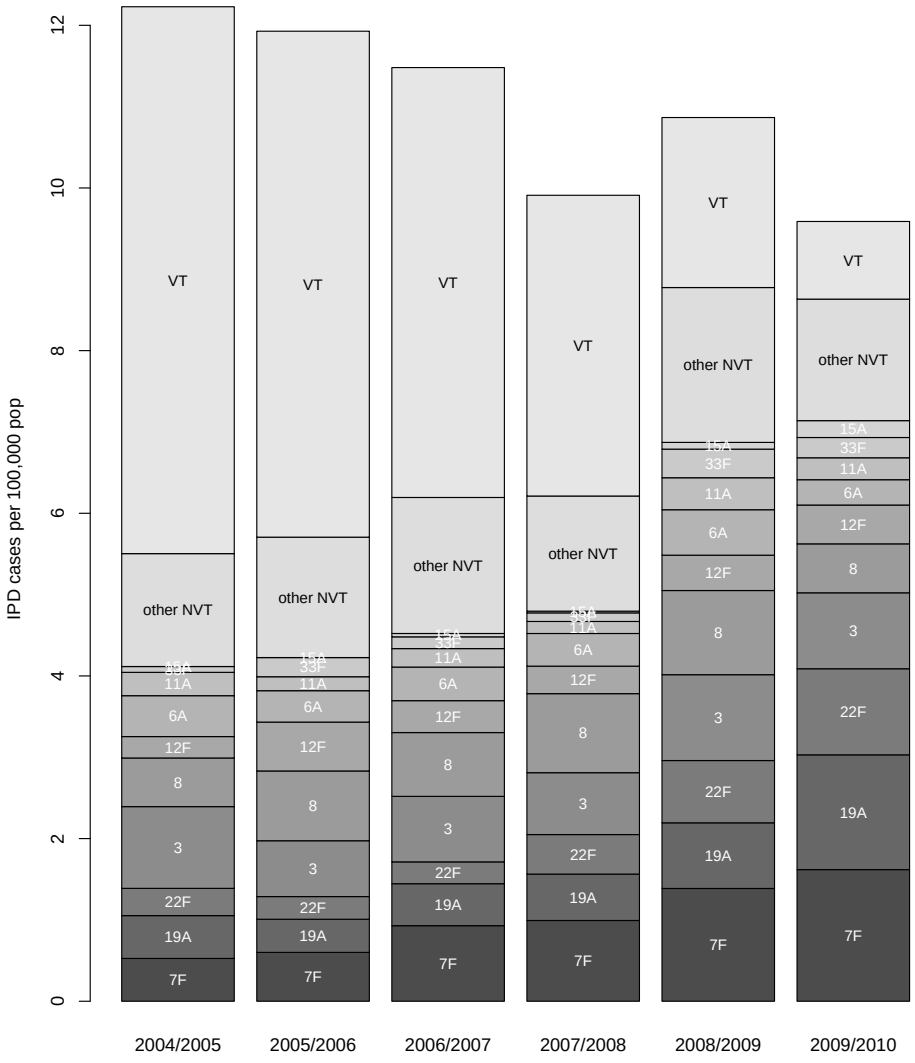


Figure 3.6: Changes in non-vaccine serotype IPD incidence distribution excluding serotype 1 for all age groups in the epidemiological years 2004/05 to 2009/10

to 64 years and over 64 years a significant reduction of vaccine type IPD of 80.0% and 85.2% with method 1 and 86.3% and 80.4% with methods 2,respectively, was observed. In the age groups <5 years and 5-64 years there was no significant increase in NVT notifications in 2008/09 compared to the predicted incidence by either method. Only in individuals aged 65 years and older a significant increase in NVT disease, 50.8% (method 1) and 46.5% (method 2), was observed. This led to no significant change in all-type incidence in this age group.

When excluding serotype 1 from the analysis serotype replacement became more apparent. Serotype one disease has been mostly affecting the age group 5-64 years. While in the over 64 year old population there was little change compared to the analysis including serotype 1, a 45% increase in NVT amongst the 5-64 year olds was estimated using method 1 although not significant in the under 5 year olds (see table 3.5), where there was a 48% increase. However, when accounting for pre- vaccination trends (method 2), no evidence for serotype replacement was detected.

The contribution of individual serotypes to the change in overall incidence can be seen in figures 3.5 and 3.6. The incidence of all serotypes included in the formulation of PCV7 decreased significantly post vaccination with most reductions being well over 70% independently of the method employed (see table 3.6 and figure 3.7). Serotypes 7F, 19A and 22F show a substantial increase following vaccination, however only the increase in the types 7F and 22F was significant when adjusting for pre vaccination trends. NVT 1 and 20 incidence substantially decreased after vaccine implementation but only the decrease in ST 1 was found significant.

The slopes of the models for VT against both the slopes for NVT including ST1 and excluding ST1 were compared. No statistical significant difference was found between NVT and VT slopes prior 2005/06 in each age group irrespective of the inclusion of serotype 1. In all age groups slopes in NVT disease were estimated to be greater then in VT disease, however excluding ST1 from the analysis caused the point estimates of VT and NVT slopes to become less distinct (see figure 3.8). This effect was most pronounced in the age group 5-64 years where most of the ST1 IPD was observed.

Discussion This analysis provides evidence that the introduction of PCV7 is associated with a substantial reduction in vaccine type IPD incidence in the

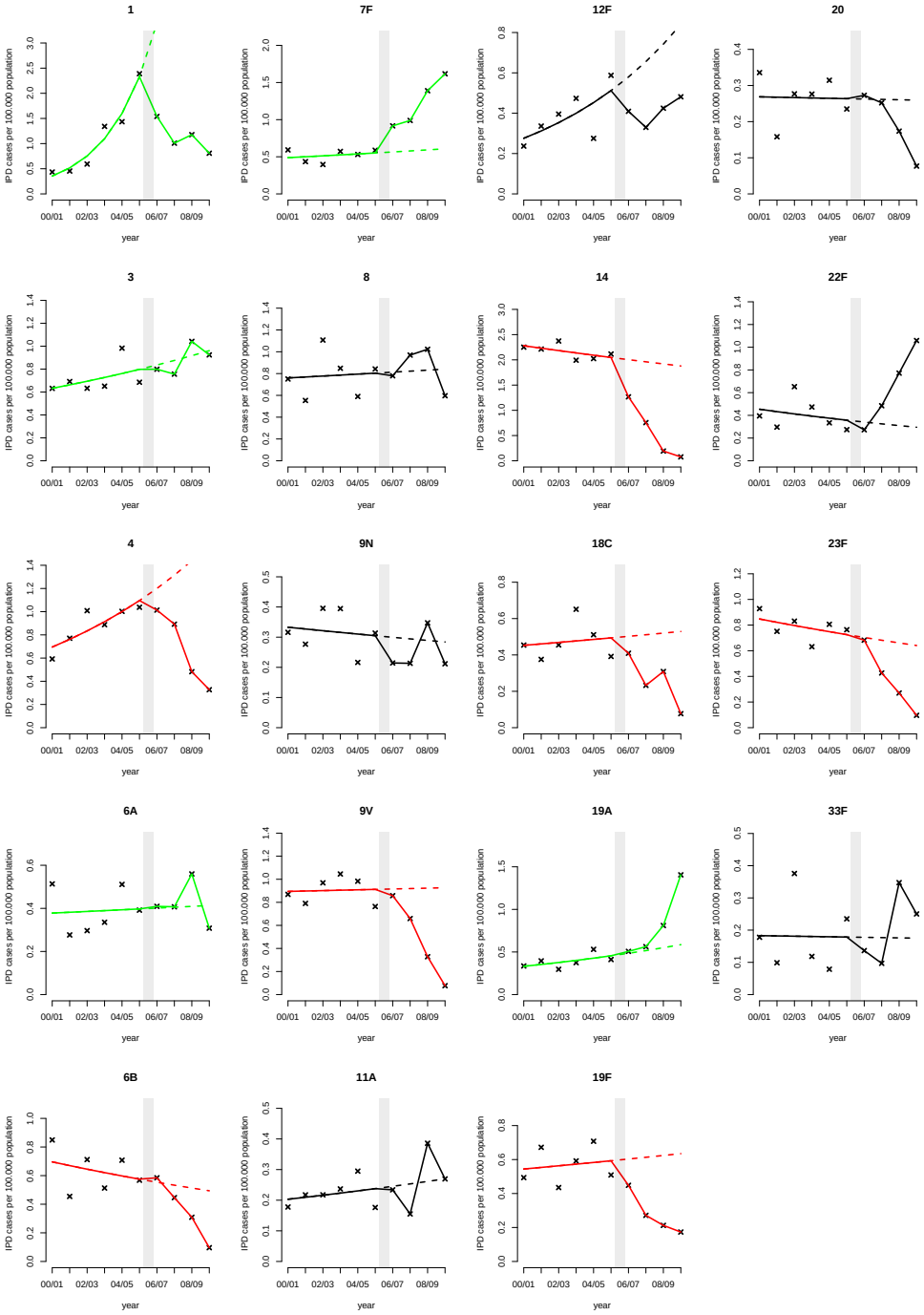


Figure 3.7: Changes in serotype-specific overall IPD incidence. The solid line indicates the model and the dashed line indicates the predicted incidence in the absence of vaccination (model 2). The serotypes included in PCV7 are marked in red and additional serotypes included in PCV13 are marked in green

Table 3.4: Changes in VT and NVT disease including serotype 1

	Incidence 2004/05	Incidence 2005/06	Incidence 2009/10	Change 2 years pre PCV7 to 2009/10 (method 1)	Change 2009/10 predicted to observed (method 2)
0-4 years					
All	38.23	27.35	14.18	-56.7% [-70.4, -40.5]	-68.5% [-80.4, -50.0]
NVT	12.87	7.12	13.49	35.0% [-12.4, 103.7]	-39.6% [-71.4, 28.6]
VT	25.36	20.24	0.69	-97.0% [-100.0, -92.0]	-97.4% [-99.6, -91.3]
5-64 years					
All	8.11	9.52	6.59	-25.2% [-35.4, -14.4]	-57.2% [-65.5, -46.9]
NVT	4.73	6.22	5.92	8.1% [-7.8, 26.2]	-45.6% [-58.3, -29.1]
VT	3.38	3.30	0.67	-80.0% [-87.2, -71.3]	-86.3% [-91.6, -78.4]
65+ years					
All	31.09	32.08	26.48	-16.2% [-28.5, -2.4]	-4.9% [-24.4, 19.5]
NVT	15.06	17.00	24.12	50.8% [25.9, 80.9]	46.5% [9.0, 97.4]
VT	16.03	15.08	2.30	-85.2% [-91.4, -77.7]	-80.4% [-88.6, -67.9]

Table 3.5: Changes in VT and NVT disease excluding serotype 1

	Incidence 2004/05	Incidence 2005/06	Incidence 2009/10	Change 2 years pre PCV7 to 2009/10 (method 1)	Change 2009/10 predicted to observed (method 2)
0-4 years					
All	35.96	24.36	11.77	-60.7% [-73.8, -44.9]	-69.8% [-81.7, -50.8]
NVT	10.60	4.12	11.07	50.4% [-6.6, 142.3]	-12.1% [-79.5, 296.8]
VT	25.36	20.24	0.69	-97.0% [-100.0, -92.0]	-97.4% [-99.6, -91.3]
5-64 years					
All	6.61	7.09	5.87	-14.3% [-26.7, -0.6]	-39.9% [-51.6, -25.4]
NVT	3.23	3.79	5.20	48.1% [23.4, 77.1]	6.4% [-13.3, 30.7]
VT	3.38	3.30	0.67	-80.0% [-87.2, -71.3]	-86.3% [-91.6, -78.4]
65+ years					
All	30.37	30.16	25.91	-13.9% [-26.6, 0.4]	1.0% [-26.4, 38.5]
NVT	14.22	14.846	23.60	62.4% [35.0, 95.2]	67.5% [2.3, 176.1]
VT	16.03	15.08	2.30	-85.2% [-91.4, -77.7]	-80.4% [-88.6, -67.9]

Table 3.6: Changes in incidence of the different serotypes

Serotype	Change 2 years pre PCV7 to 2009/10 (method 1)	Change 2009/10 predicted to observed (method 2)
PCV7 types		
4	-67.9% [-82.5, -49.1]	-79.3% [-89.0, -60.5]
6B	-84.9% [-96.9 , -68.7]	-80.5% [-93.8 , -49.0]
9V	-91.2% [-98.1 , -80.6]	-91.7% [-97.6 , -78.0]
14	-96.3% [-99.1 , -92.0]	-95.9% [-98.8 , -90.0]
18C	-82.9% [-96.4 , -60.8]	-85.5% [-96.1 , -57.0]
19F	-71.5% [-88.8 , -47.5]	-72.7% [-89.1 , -36.0]
23F	-87.7% [-97.5 , -74.8]	-85.0% [-95.1 , -62.0]
+PCV13 types		
1	-57.7% [-70.7, -42.6]	-92.3% [-95.4, -87.3]
3	10.8% [-23.4, 55.9]	-3.84% [-46.0, 73.4]
6A	-31.8% [-65.0, 16.6]	-25.5 [-69.9, 76.0]
7F	189.2% [108.8 , 310.1]	166.31% [41.1 , 412.0]
19A	198.4% [108.3 , 342.7]	139.9% [17.1 , 402.0]
other types		
8	-16.6% [-47.3 , 25.2]	-29.1% [-61.5 , 30.0]
9N	-20% [-64.9 , 57.0]	-25.5% [-72.5 , 98.8]
11A	14.4% [-45.1 , 120.6]	-0.1% [-65.5 , 194.0]
12F	11.6% [-34.6 , 80.5]	-42.7% [-74.0 , 26.0]
20	-72.0% [-94.4 , -34.6]	-70.3% [-92.8 , 3.0]
22F	248.1% [126.9 , 463.0]	258.4% [70.2 , 675.0]
33F	59.6% [-28.6 , 243.7]	42.9% [-55.3 , 372.1]

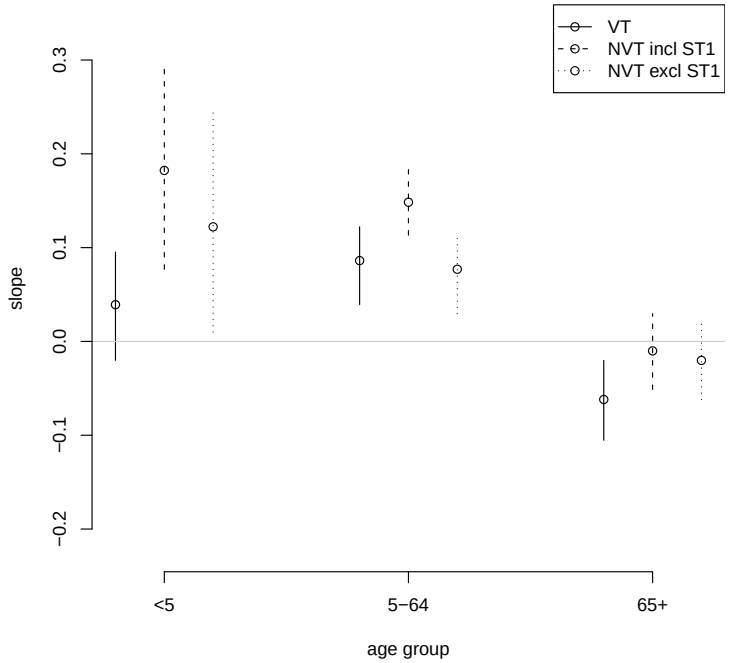


Figure 3.8: Estimated pre vaccination model slopes with confidence intervals for the VT and NTV models including and excluding ST1.

targeted age group. These effects were also experienced in older, unvaccinated population groups presumably due to induced changes in pneumococcal carriage and their consequent transmission (herd immunity). Contradicting many other studies clear evidence for serotype replacement was only found in the elderly population.

Increasing incidence in pathogens similarly depending on blood culturing as the pneumococcus were found in England and Wales amongst the under 65 year old population which might have continued after the introduction of PCV7 (see section 3.1.1). Similar trends have been observed in Scotland prior to introduction of PCV7 and, if continued, could have influenced the post vaccination analysis. Therefore, not only pre/post vaccination incidence comparison was used to assess the impact of vaccination but also a second method assuming the pre vaccination trend to continue was employed for comparison. In the absence of further information neither of the two models can be favoured and the different approaches rather represent additional uncertainty introduced by pre vaccination trends in the data. While with both models there were only small discrepancies found in the impact of vaccination on reducing VT disease, estimated NVT replacements

effects were distinct. Counter-intuitively with method 2 NVT IPD was found to have decreased post vaccination in both younger age groups. However this was found to be less pronounced when ST1 was excluded from the analysis due to its secular trends. This is, because ST1 is highly prevalent especially in the 5 to 64 year olds and decreased in prevalence after introduction of PCV7 while increasing beforehand. However, even when excluding ST1 no clear evidence for serotype replacement was found in NVT IPD post vaccination in the under 5 year olds (method 1: 50.4% [-6.6, 142.3] and method 2: -12.1% [-79.5, 296.8]) However, evidence for replacement was found in the older age groups, particularly in the individuals over 65 years of age. This is noteworthy because the replacement effects observed in the elderly population are believed to be the result of replacement occurring in the directly protected children. The apparent lack of replacement of IPD in children and adults and the strong replacement in the elderly population could be the result of age-specific differences in the invasiveness of the replacing serotypes in pneumococcal carriage. Unfortunately no pneumococcal colonisation data exists for Scotland.

Correcting for underlying trends is assuming a continuing trend of what likely is increasing ascertainment. If the pre vaccination increase in IPD is due to such, stratifying by VT and NVT should reveal similar trends. No significant difference in age-stratified pre-vaccination slopes between VT and NVT was found which suggests that these trends might have a similar source for both groups. However, when analysing type-specific changes in pneumococcal incidence an alternative approach would have been to adjust for the increase found in overall incidence rather then for the type-specific increase in order to limit the stochastic effects resulting from the low sample sizes.

No analysis on pre 2005/06 trends in pneumococcal-like bacteraemias for Scotland exists but data from England and Wales analysed in chapter 3.1.1 suggested that pre vaccination ascertainment trends in regions with a similar healthcare system to the Scottish continued in between both scenarios. The analysis was done on blood cultures only whereas this analysis consists of isolates from both blood and CSF but isolates from CSF only accounted for about 2% of all isolates included. In the population aged 65 years and older no corresponding trends were found in the control pathogens. However, in this age group there was only little pre vaccination trends in Scottish IPD notifications and extrapolating them did not change the interpretation.

In conclusion the findings provide consistent evidence that PCV7 substantially reduced the incidence of invasive pneumococcal disease associated with the targeted serotypes both directly in vaccinated individuals and indirectly through herd immunity. Further, in the 65+ population, the emergence of NVT disease almost compensated the benefits of herd immunity. In the younger age groups no significant emergence of NVTs could be detected. The sustained decrease of serotype 1 in these age groups post vaccination outbalanced the increase of the other non-vaccine serotypes, with its main drivers 7F, 19A and 22F. This reveals a general caveat in interpreting data on vaccine effects in pneumococcal disease: long-term trends in specific serotypes have been observed in the last decades [11] with the underlying driving factors being unknown.

3.1.3 England and Wales

Invasive disease

The population impact of the introduction of PCV7 to the childhood immunisation scheme on the incidence of IPD in England and Wales was recently reported [17]. Similar methods to the ones presented in section 3.1.2 were employed and trend adjusted as well as trend unadjusted estimates of the changes from pre- to post-vaccination incidence were reported on the grounds of the findings in section 3.1.1. This section will briefly summarise the main findings.

National surveillance data on invasive pneumococcal disease from July 2000 to June 2010 in England & Wales (see section 2.6) were used to calculate incidence risk ratios to calculate the impact of vaccination (see section 1.1.1). In 2009/10 following the introduction of PCV7 vaccine-type IPD was reduced by more than 90% in the under 2 year olds, over 85% in the 2-4 year olds and more than 50% in all other age groups when not accounting for underlying trends (which represents the more pessimistic scenario). Evidence for serotype replacement was found in both the very young and the older age groups; however, accounting for underlying trends (lower estimate for NVT replacement) changes in non-vaccine-type IPD were found not significant in the 5-64 year old population. The overall reduction of IPD amongst all age groups was 34% (22% when not accounting for trends).

Significant reductions of all serotypes included in the vaccine formulation reported in all age groups. The major replacing serotypes were 7F, 19A and 22F (as

found in the analysis for Scotland in the previous section). The authors also detected a pronounced long term trend of serotype 1 (upwards in 2000 - 2005 and downwards in 2006-2010) which is most likely not attributable to vaccination and could have masked parts of the serotype replacement in the 5-65 year old population; in meningitis where prevalence of ST1 is typically low the authors detected significantly increasing level of NVT disease in the 5-64 year olds.

Overall the findings are similar to those from Scotland presented in section 3.1.2. The main replacing serotypes are the same in both settings and the effect of PCV7 in reducing VT disease in both vaccinated and unvaccinated individuals was substantial. However, the apparent absence of serotype replacement amongst Scottish children remains unparalleled and the reasons for it unknown.

Carriage

This research has been published [18].

Introduction The majority of carriage episodes of the pneumococcus do not result in either local or systemic disease. It is believed that the propensity to cause disease in healthy individuals, termed invasiveness, is largely determined by the characteristics of the pneumococcus polysaccharide capsule although the explicit underlying mechanisms are yet to be identified [19, 20].

Since PCV7 is protective against invasive pneumococcal disease (IPD) [21] and carriage [22, 23], the assumption of protection of unvaccinated individuals against vaccine type (VT) IPD through herd immunity played a major role in evaluating the likely impact and cost-effectiveness of vaccination [24]. Prevention of VT carriage, however, creates a potential ecological niche in the nasopharynx for previously less prevalent serotypes to emerge (replacement). The extent to which the benefits of herd immunity will be offset by serotype replacement is hard to predict [25] and may vary by country depending on local factors such as differences in serotype distribution before vaccination and the population demography. Most surveillance systems focus on IPD and have shown large reductions in the numbers of VT cases in the targeted age groups, irrespective of vaccine schedule (see [3, 7, 26] and section 3.1.2). However differences were observed in the degree of induced herd immunity as well as in the level of non vaccine-type (NVT) replacement, the reasons for which remain unclear.

Monitoring disease outcomes provides little insight into the underlying mechanisms that determine herd immunity and serotype replacement. For a better insight on pneumococcal transmission dynamics, carriage data are essential. Carriage studies in children from Massachusetts and Norway suggest full replacement of pneumococcus in carriage after PCV7 introduction though neither study assessed the indirect effect on carriage of VT and NVT organisms in older unvaccinated age groups, nor related the changes in serotype-specific carriage prevalence to changes in IPD in the same population [2,27]. Improving the understanding of this relationship, largely determined by the invasiveness potential of the replacing NVT organisms, is essential to understanding the effect of PCV7 in different epidemiological settings.

Information on carriage in England prior to PCV7 introduction is available from a longitudinal study conducted in 2001/02 in index children and their household members [28]. Over a 10 month period pre-school children and their families were recruited by study nurses in Hertfordshire, UK, between October 2001 and July 2002. 121 families constituting 489 individuals (including 138 index children) were tested on a monthly basis for nasopharyngeal colonisation. Here the results of a cross-sectional carriage study conducted in a demographically similar population in 2008/09 are reported. The post-PCV7 findings from this chapter are compared with the pre-PCV7 baseline both for carriage and IPD to help understand the serotype-specific effects of PCV7 on carriage and thereby on IPD and predict the potential impact of higher valency conjugate vaccines on herd immunity and replacement disease.

Methods *Study population:* Children born since 4th September 2004 and thus eligible for routine or catch-up PCV were recruited along with family members from general practices in Hertfordshire and Gloucestershire. Exclusion criteria for the recruitment of index children were: moderate to severe disability, cerebral palsy, neurological disorders affecting swallowing, ear, nose and throat disorders affecting the anatomy of the ear, or immunosuppression. Local and national ethics committees approved the study protocol which was registered under NCT01040143. Written informed consent was obtained from adult study participants and from a parent/guardian of study children prior to enrolment. Information was collected on participants age, gender, household size, number of smokers in household, recent antibiotic treatment, hours in day-care and PCV7 vaccination history.

To compare to pre-vaccination carriage in England results from a longitudinal study carried out in 2001/02 in families attending the same general practices in Hertfordshire in which swabs were taken each month over a 10 month period [28] were used. At that time of the first study, serotype 6C could not be distinguished from 6A. In 2009 19 of the 122 serotype 6As from the earlier study were randomly retested, 6 of which were found to be 6C. This proportion (32%) was assumed to hold for the rest of the 6A carriage isolates from the 2001/02 study.

Specimen testing: The nasopharyngeal swabs (calcium- alginate) were taken by trained nurses and placed directly in STGG broth. Samples collected at Hertfordshire were sent by same day courier to the Respiratory and Systemic Infection Laboratory at the Centre for Infections (RSIL). They were stored overnight at 2-8°C and frozen the next morning at -80°C. Samples collected at Gloucestershire were stored locally at the Gloucester Vaccine Evaluation Unit at -80°C and transferred to RSIL in batches on dry ice. On receipt the batches were stored at -80°C. The sample then was thawed, vortexed and 50ul STGG broth was placed onto each of Columbia blood agar plate (HPA media services) with optochin disc (MAST) and streptococcus selective Columbia blood agar plate (HPA media services) and streaked out. The plates were incubated overnight at 35°C with 5% CO_2. Any colonies resembling pneumococcus were subjected to normal identification methods and serotyped using the standard laboratory protocol (compare section 2.5.3).

Statistical Analysis: Descriptive data analysis was performed in R 2.11.0 and Generalized Estimating Equations (GEE) models were analysed with STATA 10.1. Exact binomial 95% confidence intervals (CIs) were obtained for carriage rates in 2008/09 by age group (<5, 5-20,>20). To account for the longitudinal design in the 2001/02 study these carriage rates were computed using a GEE model with exchangeable correlation structure accounting for repeated sampling of individuals:

$$ln\left(\frac{p_{i,t}}{1-p_{i,t}}\right) = \alpha_0 + \gamma_i + \epsilon_{ij}.$$

Here $p_{i,t}$ is the probability of VT carriage (NVT and any carriage respectively) of individual i at time of test t, α_0 is the intercept, γ_i is the random intercept and ϵ_{ij} the error term. When comparing overall carriage as well as vaccine and non vaccine type carriage between periods this comparison took account of the longitudinal design of the 2001/02 along with other covariates by using a GEE model with exchangeable correlation structure and factors for study period, age,

gender, whether the household has a smoker and the number of children and adults in the household (compare risk factors for pneumococcal carriage in section 2.2):

$$ln\left(\frac{p_{i,t,t_p}}{1-p_{i,t}}\right) = \alpha_0 + \sum_{k=1}^{4}(\alpha_k E_k) + \psi I(t_p = T_2) + \gamma_i + \epsilon_{ij}.$$

With the notation as above here $\alpha_k E_k$ estimates the contribution of the risk factors and T_2 marks the post vaccination study period. Hence ψ estimates the change in carriage in both studies. For this analysis the data were stratified into two age groups (<5 and ≥ 5). To determine the significance of changes in carriage for individual serotypes between 2001/02 and 2008/09 Fishers exact test was used because of small numbers and statistical significance was determined accounting for multiple testing [16].

For calculating the case:carrier ratio (CCR) the invasive disease incidence of each serotype was extracted from the national surveillance database for England and Wales [4] for the epidemiological years 2001/02 and 2008/09 and related to data from the carriage studies conducted in the corresponding years. CCRs were calculated using serotype-specific carriage prevalence as numerator. Ages younger than 60 years were combined in both the IPD and carriage data sets to derive CCRs. 95% confidence intervals were calculated as described in Trotter et al [4]. For serotypes with estimates in both datasets Spearmans rank test was used to estimate the correlation of the estimated CCR's and those from by Sleeman et al. from a paediatric pre-PCV7 carriage data set in one region of England corrected for duration of carriage [29]. Since the obtained estimates for the CCRs also included adults the serotype-specific case:carrier ratio (CCR) was scaled by the average CCR from the respective study for visual comparison. For each set of estimates a serotype with a relative CCR >1 one corresponds to being more invasive than the study specific average.

To infer a rough estimate of the expected number of IPD cases after introduction of a new conjugate vaccine with increased valency a methodology following Weinberger et al. was employed [30]. For this the IPD cases in the under 60s in 2008/09, excluding the vaccine types, were inferred from the proportion of vaccine serotypes in carriage. This assumes that the vaccine types will be fully eliminated in carriage, and therefore do not cause any more disease, and full serotype replacement will happen in carriage. This approach further assumes that the ranked serotype distribution will stay the same. This assumption did not hold for the implementation of PCV7, as many new types emerged, however

Table 3.7: Overview of numbers of participants recruited, their demographic features and household (HH) structures in the 2001/02 and 2008/09 carriage studies.

	2001/02	2008/09
# Participants	489	382
# Swabs taken	3868	382
# Participants < 5 years	180	192
# Participants 5-20 years	71	57
# Participants 20+ years	237	133
# Proportion female	53.0%	56.4%
# HH	121	146
Median HH size (range)	4 (2-7)	4 (3-7)
Median # adults in HH (range)	2 (1-5)	2 (1-5)
Median # children in HH (range)	1 (1-3)	2 (1-4)
Proportion of smoke free HH	66.9%	81.0%

the ranking of serotypes might be pre-defined to some extend by the capsular expression [20]. PCV10 and PCV13 will cause less perturbation to carriage so this assumption might be reasonable.

Simpsons index for diversity was calculated to assess the potential change in diversity in the bacterial population due to vaccination [31]. The ranked serotype distribution was compared to the pre-vaccination distribution and confidence intervals were obtained using the methods described by Hanage et al. [32]. To ensure that only a single isolate per carriage episode was included consecutive swabs with the same serotype in the 2001/02 study were excluded on the assumption that it was carriage persisting from the previous month.

Results 400 individuals were enrolled between 24/04/2008 and 09/11/2009. One participant withdrew before being swabbed and in 17 individuals swabbing had to be aborted early; these 18 were excluded from further analyses. The demographic features of the remaining 382 were similar to the participants in the 2001/2 study apart from the proportion of households with at least one smoker which was lower in the more recent study (Table 3.7). Of 180 children eligible for catch-up or infant vaccination only 4 were unvaccinated. The serotype specific carriage rates in comparison with the 2001/02 study are shown in table 3.8.

A pneumococcus was grown from 127 of the 382 (33.2%) swabs and a serotype determined in 123 (97%). The most prevalent serotypes were 19A (10), 23B (9), 11C (8), 15B (8), 21 (8), and 6C (8). Compared to pre-vaccination levels, there was a significant reduction ($p<0.007$) in carriage of VTs 6B, 14, 19F and 23F (odds: 0.13 [0, 0.36], 0 [0, 0.32], 0.23 [0, 0.62], 0.08 [0, 0.4] respectively). Also

Table 3.8: Number of isolates found in carriage in 2001/2002 and 2008/2009. *6A and 6C were not distinguished in 2001/2002.

	2008/09				2001/02			
	<5y	5-20y	≥20y	All	<5y	5-20y	≥20y	All
10A	2	1	1	4	1	0	1	2
11A	4	0	0	4	27	8	3	38
11B	3	0	1	4	0	0	0	0
11C	6	1	1	8	0	0	0	0
14	0	0	0	0	64	24	5	93
15B	6	0	2	8	5	0	0	5
15C	4	0	1	5	14	0	0	14
16A	0	1	0	1	0	0	0	0
16F	0	0	0	0	3	1	9	13
17A	1	0	0	1	0	0	0	0
17F	0	0	0	0	0	0	2	2
18B	0	0	0	0	2	0	0	2
18C	2	1	0	3	13	3	9	25
19A	10	0	0	10	18	3	2	23
19D	0	0	0	0	1	0	0	1
19F	3	0	0	3	100	24	3	127
20	0	0	0	0	0	1	1	2
21	5	2	1	8	6	0	4	10
22A	1	0	0	1	0	0	0	0
22F	3	1	0	4	14	7	2	23
23A	4	1	0	5	9	5	0	14
23B	6	0	3	9	0	0	0	0
23F	1	0	0	1	97	10	13	120
24F	2	0	0	2	0	0	0	0
27	0	0	0	0	1	3	3	7
28F	1	0	0	1	0	0	0	0
29	1	0	0	1	0	0	0	0
3	5	0	1	6	3	7	0	10
31	2	0	0	2	1	1	0	2
33A	2	0	0	2	0	0	0	0
33F	4	0	0	4	1	0	0	1
34	0	1	1	2	0	0	1	1
35A	0	0	0	0	0	0	1	1
35B	0	0	0	0	4	2	1	7
35C	0	0	0	0	0	0	1	1
35F	1	0	0	1	16	0	0	16
37	1	0	0	1	0	0	5	5
38	0	0	1	1	1	0	0	1
4	0	0	0	0	0	0	3	3
6A*	2	0	0	2	94	21	7	122
6B	2	2	0	4	193	11	10	214
6C*	8	0	0	8	0	0	0	0
7F	2	2	2	6	2	0	0	2
8	0	0	0	0	0	2	0	2
9N	0	0	0	0	5	2	0	7
9V	1	0	0	1	12	7	4	23
NTR	3	0	1	4	20	5	8	33

a non significant (p=0.014) reduction of carriage of the vaccine associated type 6A was found (odds: 0.24 [0, 0.78]). For the remaining PCV7 types no carriage episodes of serotypes 4 and 9V were found post-vaccination, but pre-vaccination levels were too low to detect any significant change (odds: 0 [0, 17.38], 0 [0, 1.41]). VT 18C was identified in three out of 382 (0.79%) swabs in 2008/2009 and in 25 out of 3,868 (0.64%) in the 2001/2002 study (odds: 1.22 [0, 3.48]). NVTs 33F, 7F, 10A, 15B, 21 and 19A significantly increased ($p<0.0012$) in carriage with odds of 40.9 [5.3, Inf], 30.8 [6.8, Inf], 20.4 [3.8, Inf], 16.5 [5.6, Inf], 8.2 [3.3, Inf] and 4.5 [2.1, Inf] respectively. A significant increase was also found in serotypes 23B, 11C and 11B which were only detected in the post-vaccination study.

The proportion of swabs with VT and NVT serotypes according to age-group in both studies is shown in table 3.9. The factors smoking and the number of children or adults in the household were not significant in most models for assessing the change in carriage post vaccination and therefore excluded from the analysis. The odds ratio of overall VT carriage post-vaccination compared to pre-vaccination using the GEE with binary outcome was estimated to be 0.10 95%CI (0.05-0.20). For NVT carriage the estimate is 3.93 95%CI (2.97-5.22) and no significant effect on overall carriage was detected: 1.18 95%CI (0.91-1.54) (Table 3.10). When applying the same models to individuals younger than five years only or individuals more than 5 years of age similar patterns were found.

Inferring the change in IPD from 2008/09 by predicting changes in pneumococcal carriage after introduction of the conjugate vaccines is was estimated that continuing with PCV7 will eventually reduce the burden of IPD a further 7%. PCV10 will reduce IPD in the under 60 years olds by 39% and PCV13 by 49%. For comparison: the vaccine coverage among the 2008/09 isolates of IPD for PCV7, 10, 13 was 15.2%, 57.8%, 77.6% respectively.

Simpsons index of diversity for the 2001/02 samples was 0.908 95%CI: (0.899, 0.917) and increased significantly in the 2008/09 samples to: 0.961 95%CI: (0.953, 0.969). Furthermore, the ranked frequency distribution of the serotypes changed to become more distinct from the distribution found pre vaccination (Figure 3.10) which was similar to the one observed by Hanage et al. in the pre-vaccination period [32].

Prior to its introduction PCV7 included types responsible for similar proportions of carriage episodes (62.2%) and disease (55.9%). In 2008/9 the additional types covered by higher valency vaccines were more prevalent in IPD than carriage,

Table 3.9: Number and proportion of positive VT, NVT and All (including non-typeable) carriage isolates in 2008/09. $^{+}$The proportion for 2001/02 was calculated accounting for multiple testing of the participants.

		NVT	VT	ALL
<5 years	Cases 08/09 (n=192)	87	7	98
	Proportion 08/09	45.3% (38.5-52.6)	3.6% (1.0-6.2)	51.0% (43.8-58.3)
	Proportion 01/02^{+}	15.3% (12.7-18.3)	31.9% (28.1-36.1)	48.4% (44.1-52.7)
5-20 years	Cases 08/09 (n=57)	15	0	16
	Proportion 08/09	26.3% (15.8-38.6)	0% (0-6.4)	28.1% (17.5-40.4)
	Proportion 01/02^{+}	9.1% (6.3-12.8)	9.9% (7.3-13.3)	20.6% (16.1-26.1)
>20 years	Cases 08/09 (n=133)	10	3	13
	Proportion 08/09	7.5% (3-12)	2.3% (0-5.3)	9.8% (5.3-15)
	Proportion 01/02^{+}	3.3% (2.4-4.8)	4.1% (3.0-5.5)	7.6% (6.2-9.5)

Table 3.10: Odds ratios for comparing 2001/02 to 2008/09 carriage using GEE to account for multiple testing in the earlier study. Fixed effects were the study period (outcome measure) and antibiotic treatment.

	<5	>5	ALL
VT	0.09 (0.04-0.20)	0.26 (0.08-0.83)	0.10 (0.05-0.20)
NVT	4.52 (3.11-6.60)	2.94 (1.79-4.85)	3.93 (2.97-5.22)
ALL	1.22 (0.870-1.74)	1.41 (0.90-2.21)	1.18 (0.91-1.54)

particularly the additional three in PCV10 which comprised 32.6% of IPD but only 4.7% of carried isolates (Table 3.11).

The ranking of carried serotypes by frequency of detection in the post-PCV7 dataset and their associated CCRs as estimated from 2008/9 carriage prevalence data are shown in Figure 3.9. The data are provided in table 3.12. The CCR estimates of Sleeman et al. estimated from carriage incidence data [29] are also presented in Figure 3.9. CCR estimates from these two methods were highly correlated ($p<0.001$, $\rho=0.72$) and the rankings still distinguished the more from the less invasive serotypes. From the 15 most prevalent serotypes in carriage in 2008/09 19A, 3, 7F and 22F stand out with a generally higher CCR. Despite their high incidence in invasive disease serotypes 1, 8, 12F, 4 and 14 (1.14, 0.58, 0.25, 0.22, 0.14 cases per 100,000 population respectively in under 60 year olds) were not detected in carriage. On the other hand despite being found in 2008/09 carriage serotypes 11C, 16A, 17A, 28F and 33A were not found in 2008/09 IPD at all and only caused 0, 1, 0, 1 and 0 cases respectively of invasive disease out of over 13,000 isolates serotyped between July 2006 and June 2009.

Discussion This study documents the changes in carried pneumococci following the introduction of PCV7 in England and relates these to concomitant changes in disease in order to assess the invasive potential of the serotypes now predominating carriage. This knowledge is essential for understanding replacement disease and provides insight into the likely population impact of higher valency vaccines.

As reported elsewhere [2, 27], a major reduction in VT carriage in vaccinated children under 5 years was found, but no overall change in carriage prevalence due to replacement with NVTs. In contrast, there was an overall reduction in IPD in this age group (see section 3.1.2 and 3.1.3), illustrating that the outcome of the PCV programme as expressed in IPD is determined by the invasiveness potential of the individual NVTs emerging in carriage. Overall carriage prevalence in older

Table 3.11: Number and percentage of pneumococcal isolates included in the different vaccine formulations in those aged less than 60 years caused by serotypes included in PCV7, in PCV10 and not in PCV7, in PCV13 and not in PCV10 and the remaining serotypes. Confidence intervals were derived from a multinomial distribution

	2008/09		2001/02	
	Carriage (%)	% in IPD	Carriage (%)	% in IPD
PCV7	11 (8.7 [3.9, 13.4])	15.2 [13.9, 16.5]	605 (62.2 [59.2, 65.3])	55.9 [53.8, 58.1]
+PCV10	6 (4.7 [1.6, 8.7])	32.6 [30.9, 34.4]	2 (0.2 [0.0, 0.5])	10.2 [8.9, 11.5]
+PCV13	18 (14.2 [8.7, 20.5])	15.9 [14.6, 17.2]	155 (15.9 [13.6, 18.3])	8.9 [7.7, 10.2]
Rest	92 (72.4 [64.6, 79.5])	36.3 [34.5, 38.1]	210 (21.6 [19.0, 24.3])	25.0 [23.1, 26.9]

Table 3.12: Estimated CCR for all serotypes found in carriage in 2008/2009 with the corresponding number of isolates found in carriage and in IPD in the under 60y-old population. Some serotypes were found in IPD but not in carriage: sertoype 1 (387 isolates), 8 (196), 12F (83), 4 (74), 9V (58), 14 (48), 5 (23), 20 (21), 15A (12), 17F (10), 16F (10), 35B (6), 27 (6), 13 (4), 28A (3), 12B (3), and one each of 9L, 7C, 7B, 7A, 7, 6, 35A, 28, 18A, 10F. *A total of 81 isolates of 6A/6C were found in IPD of which one-third was assumed to be 6C and the rest 6A.

	Carriage 2008/09	IPD 2008/09	Cases per 100,000 carriers
19A	10	170	18.6 (10.3 - 33.7)
23B	9	11	1.3 (0.7 - 2.5)
11C	8	0	0 (0 - 0)
15B	8	24	3.3 (1.7 - 6.4)
21	8	7	1 (0.5 - 1.9)
6C*	8		3.7 (1.9 - 7.2)
3	6	134	24.4 (11.5 - 52.5)
7F	6	326	59.3 (27.9 - 127.6)
15C	5	29	6.3 (2.8 - 14.6)
23A	5	14	3.1 (1.3 - 7.1)
10A	4	27	7.4 (2.9 - 18.7)
11A	4	30	8.2 (3.3 - 20.7)
11B	4	1	0.3 (0.1 - 0.7)
22F	4	157	42.9 (17.1 - 108.5)
33F	4	49	13.4 (5.3 - 33.9)
6B	4	38	10.4 (4.1 - 26.3)
18C	3	38	13.8 (4.8 - 40)
19F	3	45	16.4 (5.7 - 47.4)
24F	2	6	3.3 (0.9 - 11.7)
31	2	14	7.6 (2.2 - 27.4)
33A	2	0	0 (0 - NaN)
34	2	4	2.2 (0.6 - 7.8)
6A*	2		29.5 (8.3 - 105.5)
16A	1	0	0 (0 - NaN)
17A	1	0	0 (0 - NaN)
22A	1	2	2.2 (0.4 - 12.1)
23F	1	42	45.9 (8.4 - 254)
28F	1	0	0 (0 - NaN)
29	1	1	1.1 (0.2 - 6)
35F	1	10	10.9 (2 - 60.5)
37	1	2	2.2 (0.4 - 12.1)
38	1	9	9.8 (1.8 - 54.4)
9N	1	32	34.9 (6.4 - 193.6)
Not typed	4	546	

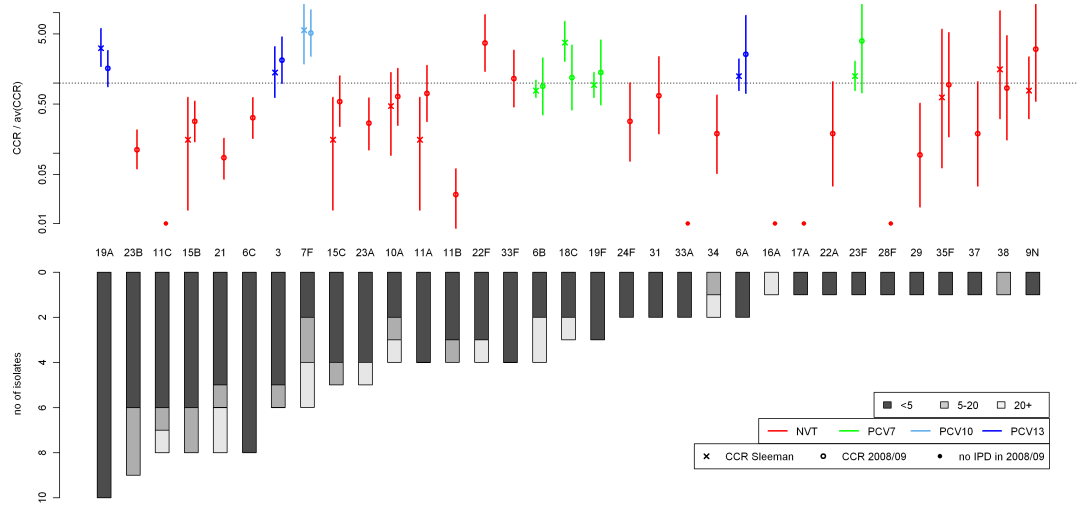


Figure 3.9: Serotype distribution in carriage in 2008/09 (below) and scaled CCR by Sleeman et al. and from the 2008/09 data (above). *6A and 6C not distinguished in Sleemann et al. Where no serotype data from Sleemann et al. was available only the estimate obtained from E&W data is shown.

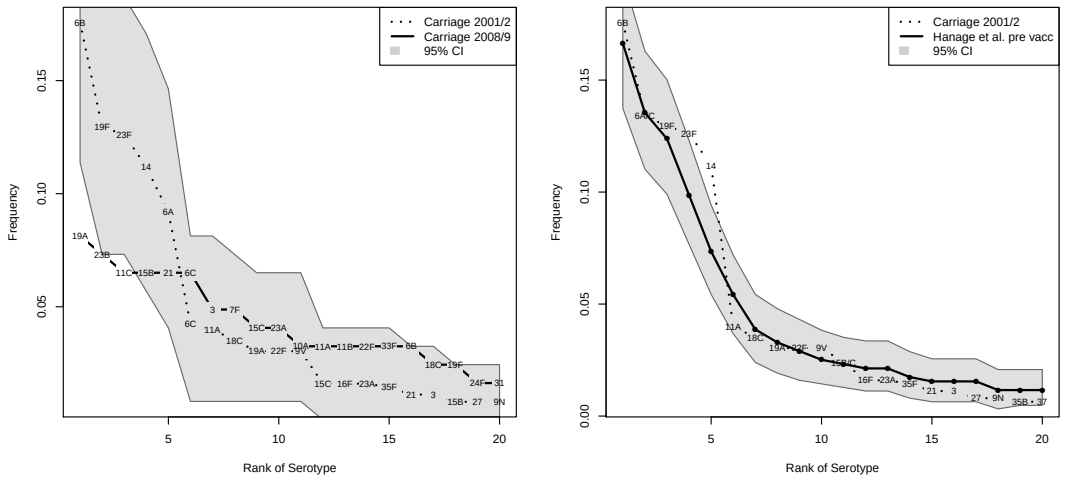


Figure 3.10: Changes in ranked serotype distribution from 2001/02 to 2008/09 with 95% confidence bounds on the rank (top) and from the pre-vaccination distribution by Hanage et al. to the findings from 2001/02 (bottom). For comparison with the findings with Hanage et al. 6A and 6C was aggregated to 6A/C and 15B and 15C to 15B/C in the bottom graph.

unvaccinated siblings and parents was somewhat higher post-PCV7 due to a large increase in NVT and a smaller non-significant reduction in VT carriage. However, IPD in these older age groups has not shown an overall increase in the UK [17], indicative of the lower overall invasiveness of the replacing NVTs.

This study shows that PCV7 provided protection against serotypes that were highly prevalent in both disease and carriage in the UK. The additional serotypes covered by PCV10 and PCV13 are now responsible for a large proportion of invasive disease, but were found relatively rarely in carriage (Table 3.11). While further replacement in pneumococcal carriage is likely to occur after introduction of these higher valency vaccines, the findings suggest that since most of the potential replacement types identified have lower CCRs they will cause less invasive disease. However, serotypes like 22F and especially the ones not found in carriage but present in IPD (e.g. serotype 8 and 12F) could reduce the overall benefits of higher valency vaccines. Interestingly, the three additional serotypes covered by PCV10 had a very low carriage prevalence accounting for <5% of the carried serotypes in 2008/9 but >30% of IPD cases whereas the further three in PCV13 are more similar to the PCV7 types, being similarly prevalent in carriage and disease. While changing to PCV10 has therefore less potential to prevent IPD than PCV13 it may cause fewer perturbations in the nasopharyngeal pneumococcal population. Comparative carriage studies in countries using PCV10 with those using PCV13, or with different PCV coverage of prevalent serotypes before introduction, would be informative to help understanding the carriage dynamics underlying serotype replacement.

The diversity of the population of pneumococcal serotypes amongst carriage isolates in the absence of any external pressure is thought to be relatively stable [32]. If this population is challenged by vaccination with a reduction in the dominance of a few highly prevalent types the diversity increases and the population takes time to return to the previous level of diversity. Hanage et al. suggested methods of assessing these changes: Simpsons Index of diversity and the concept of a typical distribution for the ranked frequency of the serotypes. Applying these to pre and post-vaccination carriage data, a significant increase in diversity from 0.908 in 2001/02 to 0.961 in 2008/09 was found, consistent with the PCV7-induced changes in the bacterial population still evolving at that time. Evidence for this can also be found in the ongoing changes in non-PCV7 IPD in 2009/10, prior to introduction of PCV13. These show a continuing increase in the 6 additional serotypes covered by PCV13 but a decrease in non-PCV13 serotypes in children

under 2 years compared with 2008/9 [33]. With the introduction of PCV13 in the UK in March 2010 [34] the longer term impact of PCV7 on carriage and IPD will not be evaluable though its effect may still be operative and not all future changes necessarily attributable to PCV13.

Recently developed molecular serotyping methods found up to nine times higher proportions of multiple carriage than detectable with standard WHO culturing methods [35]. Using the WHO method 1 (0.26%) multiple carriage episode in 2008/2009 and 4 (0.10%) in 2001/02 were identified. Undetected episodes of multiple carriage would result in over estimation of CCRs. However, direct comparison of molecular with conventional serotyping methods has so far only been performed on specimens from developing countries where carriage prevalence is very high [36,37]. In such settings the molecular method might reveal more multiple carriage episodes than in countries such as England where carriage prevalence is lower. Furthermore, detecting multiple serotype carriage is likely to primarily uncover carriage episodes of serotypes previously found to be less prevalent [38]. Therefore the potential bias introduced by the WHO standard culturing methods is believed to have little impact on the inferences from the CCR since the focus is on the types more common in carriage.

This study has some limitations. First, the earlier study had a longitudinal design while the recent study was cross-sectional. However, the estimates accounted for multiple sampling of individuals in the earlier study as well as differences in age distribution within the age groups, gender, exposure to smoke, and household size by using a GEE, which is designed to fit the parameters of a generalised linear model in the presence of unknown correlation. Second, owing to the lack of power of serotype-specific carriage data in adults, data of children and adults were pooled to derive the CCR, despite different age distributions in the samples for IPD and carriage. Previously reported CCR estimates for children and adults in England and Wales [4] using the carriage data from the earlier study are highly correlated (see figure 3.11), supporting the use of pooled carriage data from children and adults in the later study. Third, secular changes in serotype distribution in IPD can occur in the absence of vaccination [12], which may be due to alterations in carriage prevalence. With the cross-sectional design of the 2008/ 2009 study, this could not be accounted for. However, in England the only major secular change in the serotypes causing IPD observed over the last decade has been in serotype 1 [17], which was not detected in either the pre- or post-PCV7 carriage studies. Fourth, invasion is thought to follow shortly after

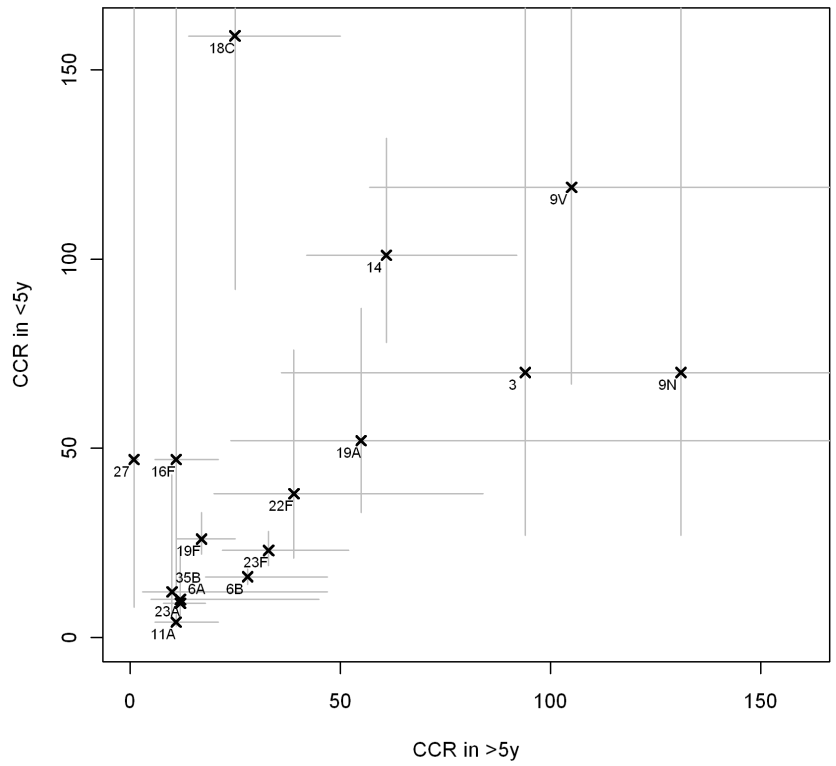


Figure 3.11: Estimated CCR in children and adults from Trotter and colleagues [4]. The grey lines represent the confidence bounds. Spearman's rank test for correlation: p=0.01, ρ =0.62.

acquisition of carriage rather than being a constant risk throughout the duration of carriage [39]. Thus, a further potential limitation of this study is that CCRs were estimated using carriage prevalence rather than the incidence of new carriage episodes, the latter being derived using prevalence and carriage duration.

Few data on serotype-specific duration of carriage are published, and for the serotypes newly emerging after introduction of PCV7, no information is available. Therefore, carriage prevalence was used to get an estimate of the CCRs. Where information on CCRs estimated using carriage incidence was available [29], a high correlation with the estimates derived here was found. Furthermore, the estimates for the CCRs in this study were consistent with those derived from 2001/2002 carriage and IPD (see figure 3.12), showing that this measure is stable over time. Hence, the current estimates of the CCR in this thesis should be sufficient to distinguish serotypes with lower invasiveness from those with higher invasiveness.

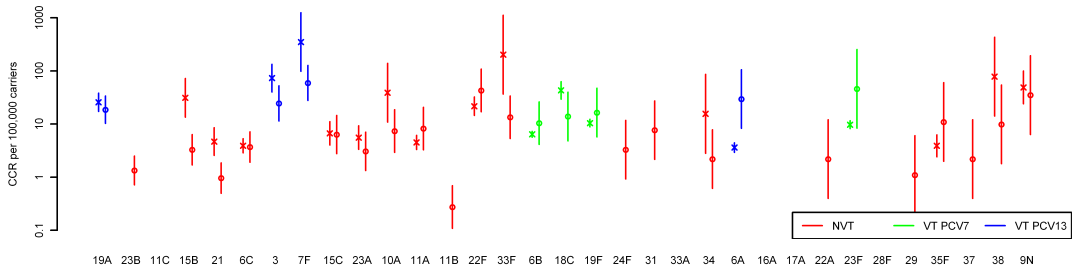


Figure 3.12: Serotype-specific CCR estimated from carriage in disease in 2001/02 (crosses) and 2008/09 (dots).

In conclusion, this study illustrates the value of generating carriage data in parallel with IPD surveillance data to help understand the serotype-specific changes in IPD observed in different epidemiological settings and predict the effect of higher valency vaccines. Evidence is provided that the incremental benefit on IPD of the recent switch from PCV7 to PCV13 in the UK is likely to be substantial but may be offset by further increases in serotypes 8, 12F and 22F. Such emerging serotypes with high CCRs are potential candidates for inclusion in future conjugate vaccines. More research to elucidate the serotype-specific capsular properties [20, 40] or other factors associated with carriage and invasiveness is needed in order to understand better the likely impact of future conjugate vaccines.

3.2 The pneumococcus outside the UK

3.2.1 The United States of America

In the United States PCV7 was introduced into the national childhood immunisation scheme in a in a four dose (3+1 booster) shedule in the second half of 2000 including a catch up campaign for children at risk aged 2-4 years. By August 2001, a shortage of vaccine supply was reported [41] which is likely to have introduced a temporary decrease in vaccine uptake. However vaccine uptake increased from 89% ($\leq$ 1 dose PCV7) and 68% ($\leq$ 3 doses PCV7) among children born in 2001 to 95% and 84% respectively among children born in 2005 [42] and is assumed to be stable since then. The use of PCV13 was recommended by the Advisory Committee on Immunization Practices in March 2010 to replace PCV7 for all children and a supplemental dose for those between 14 and 59 month was advocated. The transition was was fast (after two month 50% of doses administered were PCV13 instead of PCV7) and uptake levels remained high [43].

In the US, data on pneumococcal carriage, prior to the introduction of PCV7 to the national immunisation scheme, was only available in children under two years of age presenting with acute otitis media [44]. Therefore a study conducted in 2001, just after introduction of the vaccine is often used as a proxy for colonisation before introduction of PCV7 [45]. From specimens obtained in children younger than 7 years of age, presenting for well-child or sick visits to their primary care practice in Massachusetts, Huang et al. found a decrease in VT colonisation from 36% to 14% from 2001 to 2004. This was fully balanced by a concomitant increase in NVT carriage so that overall colonisation rates stayed the same. In a follow up the authors reported a further decrease in VT and increase of NVT carriage [2]. By studying the change in diversity of the bacterial population after vaccination, Hanage et al. suggested that serotype replacement following the introduction of PCV7 in the US (at least in Massachusetts) was complete in 2009 [32]. No information on changes to nasopharyngeal colonisation rates introduced by the transition to PCV13 are available yet.

The Active Bacterial Core surveillance (ABCs) is an active population- and laboratory-based system which provides data on invasive pneumococcal disease for eight states in the US. There is some variation in the number of reporting sites over time. Evaluating this data, even during the period of shortage in vaccine

supplies, a high impact on invasive disease among all age groups was reported [1] and found to continue further [3, 46]. Vaccine types did practically vanish by 2007 when comparing IPD incidence to the pre vaccination average from 1998-99 showing a marked reduction of 94%. At the same time NVT invasive disease (excluding serotype 19A) increased by 30% with an outstanding increase of IPD attributable to serotype 19A of more then 300%. However, overall IPD was found to have declined by 45% which was significant for all age groups [3].

Marked reductions were also reported in meningitis and non invasive pneumococcal associated disease endpoints. Hsu et al. [47] found a 30% decline in overall pneumococcal meningitis. Steenhoff et al. [48] suggested that the overall incidence of pneumococcal bacteraemia declined by 57% until 2005. Eskola et al. reported a 34% decrease in pneumococcal associated otitis media [49]. Grijalva et al. found evidence for a 35% decline in all-cause pneumonia hospitalisation in children up to two years of age [50] and Tsai et al. [51] for a 66% decline in hospitalisations for pneumococcal meningitis in the same age group.

In general a consistent significant reduction in pneumococcal disease reports among all age groups was reported. Also some evidence for NVT replacement was found and serotype 19A in particular emerged.

3.2.2 The rest of the world

The early findings from the US already suggested that the introduction of PCV7 not only reduced the burden of disease but also was highly cost effective [24]. So subsequently most of Europe [52], Canada, Mexico, Australia and New Zealand as well as some developing countries introduced a conjugate pneumococcal vaccine (mostly PCV7) to their national childhood immunisation scheme. The Global Alliance for Vaccines and Immunisation (GAVI) has helped with the funding for a wider distribution of the pneumococcal conjugate vaccines amongst developing countries [53]. Surveillance systems, mostly focussing on invasive disease, have generally shown large reductions in VT incidence in the targeted age groups [3, 7, 26, 54]. However differences were reported regarding the degree of induced herd immunity as well as the level of NVT replacement. The potential reasons for this could be manifold. They include:

Schedule PCV7 was licensed in a 3+1 schedule (3 doses and additional booster dose). Some variations have been studied and implemented. However, the

2+1 schedule implemented e.g. in the UK these was found to induce similar immunogenic response [55] and was found similarly effective to the complete schedule [56]. A 3+0 schedule was implemented in Australia and efficiently reduced VT disease [57].

Serotype distribution Pre vaccination serotype distribution varies to a great extend worldwide [58, 59]. The impact of introducing PCV is likely to be a function the proportion of VT carriage pre vaccination which might have lead to difference in the observed post-vaccination epidemiology.

Vaccine coverage / catch up The coverage of the vaccine amongst the children eligible for vaccination can be affected by things like the general attitude of a population towards vaccination. Reimbursement strategies, which vary across different countries, are likely to play a major role for the compliance of the population. In some countries a catch up was set in place to rapidly increase the vaccine coverage whereas in others the coverage was reduced through a shortage in vaccine supplies (for an overview see table 3.13).

Timing When comparing the vaccine effects of different countries one should also consider the time since implementation of the program. The full effect of the vaccine program is likely to only become apparent after a couple of years. Hanage et al. proposed, that in the US, the indirect effects of vaccination in children stabilised seven years after introduction of PCV7, yet were ongoing after four years [32]. Section 3.1.3 showed that the epidemiology in England and Wales is still likely to change more than two years after introduction of PCV7 even though a catch-up campaign initiated high uptake and no supply shortage was reported [18].

Population structure Population demographics induce differences in the transmission dynamics of the pneumococcus. Even though differences between industrialised countries are probably negligible, analysis of developing countries should take into account the major difference in age distribution and mixing patterns as well as a rapidly changing age distribution through increasing life expectancy.

Trends As outlined in section 3.1.1 pre vaccination trends could indicate changes in ascertainment due to numerous reasons including increasing blood culturing practice and an increasing number of laboratories choosing to report. Also increasing automation of detection techniques or improved survival of

people with underlying conditions, which could have increased the numbers of vulnerable people in the population, could have contributed. Additionally secular trends as reported by Harboe et al. [12] and also found for the UK (see Miller et al. [17] and sections 3.1.2 and 3.1.3) could alter the interpretation of the data.

Reporting Differences in incidence levels do not necessarily reflect different disease presence. Reporting behaviour including the catchment area of the surveillance system, identification rates and the level of voluntary reporting of non-notifiable disease outcomes could be dissimilar. However, if these were relatively stable over time reporting behaviour should not influence the assessment of change in incidence post vaccination.

3.2.3 Post PCV7 changes in comparison

An overview of the differences in timing, schedule and catch up for the introduction of PCV7 for many countries which introduced PCV7 is provided in table 3.13. To address further differences in reporting and timing between countries with sufficient post vaccination IPD data the following paragraphs will discuss similarities and differences in the experience of the US, Australia and the UK. This discussion follows a presentation held by Elizabeth Miller at the 7th International Symposium on Pneumococci and Pneumococcal Diseases (ISPPD-7) [60]. Since the observed direct impacts were rather similar this comparison focussed on the replacement effects induced by vaccination and is split into separate analysis for the groups with the highest burden of pneumococcal disease: children and elderly. Data on IPD in England and Wales was extracted from the enhanced national surveillance (see section 2.6.1). Published data on IPD incidence for the US [3] was also included and so was unpublished data from the Australian routine surveillance for IPD [60]. In all countries a different vaccination schedule was implemented (3+1 in the US, 2+1 in the UK and 3+0 in Australia). Vaccine coverage in all three settings was reported to reach high levels soon after initialisation of the program. In all three settings a substantial reduction in VT disease was found but here the focus is on differences observed in NVT replacement disease. A more in depth review has been initiated by an World Health Organisation (WHO) expert meeting in July 2010. First results became available at the WHO Strategic Advisory Group of Experts (SAGE) in 2011 [8, 61].

Table 3.13: Overview of the introduction of the 7-valent conjugate vaccine into universal childhood vaccination schemes, the proposed vaccination schedule and whether a catch up campaign was in place to rapidly increase the uptake.

	Introduction	Scheme	Catch up	Notes
Australia	2005	3+0		
Austria	September 2004	3+1	No	Only reimbursement for risk groups
Belgium	January 2005	2+1	Yes	
Canada	Early 2002	3+1	No	Catch up in high risk <5y
Cyprus	August 2008	3+1	Yes	
Denmark	October 2007	2+1	Yes	
France	June 2006	2+1	Yes	Cost sharing/ full reimbursement[a]
Germany	July 2006	3+1	Yes	
Greece	March 2006	3+1	Yes	
Hungary	October 2008	3+1	Yes	
Ireland	September 2008	2+1		
Italy	May 2005	2+1	No	Not in all regions / regional reimbursement variations
Korea	November 2003		No	No universal introduction but about 78% newborns in 2006 got vaccinated on private costs
Luxembourg	October 2004	3+1	Yes	
Netherlands	June 2006	3+1		
Norway	July 2006	2+1	Yes	
Slovakia	April 2008	2+1		Cost sharing[b]
Sweden	January 2009	2+1		
Switzerland	November 2005	2+1	Yes	At risk introduction in July 2001 at 3+1
UK	September 2006	2+1	Yes	
USA	October 2000	3+1	No	

[a] 80% of the population has a private insurance which covers the rest of the costs

[b] Full cost coverage only for risk groups

Children In children younger than the age of 5 living in England or Wales NVT incidence per 100,000 population rose from 7.8 in the pre PCV7 era (2004/05 & 2005/06 average) to 14.5 in 2008/09 representing a 86% increase (less if corrected for pre vaccination trends). Similar levels were reported from the Australian program which was implemented earlier: in the pre vaccination era between 2002 and 2004 the average incidence per 100,000 population was 7.2 and rose to 15.3 in 2007/08 reflecting a 113% increase [60]. In the US, even though PCV7 was introduced much earlier, much less replacement was reported. From pre PCV7 levels of 16.8 cases per 100,000 population on average between 1998 and 1999 notification of NVT IPD increased to 22.1 (32% increase). However, cases of IPD in England and Wales as well as in Australia represent severe and therefore hospitalised cases whereas in the US many outpatients are included as well which might account for the higher incidence (pre-vaccination) reported compared with England and Wales and Australia. Data on hospitalised cases of NVT IPD in the US were more in line with what was observed in England and Wales and Australia: pre vaccination levels of 6.1 in 1998/99 increased by 102% to 12.3 cases per 100,000 population in 2006/07. These findings are similar to the ones from Canada [62] and the Netherlands [7], however, other studies reported different findings [56, 63].

Elderly Pre vaccine trends due to increased ascertainment found no support in the elderly as analysed in section 3.1.1. In individuals older than 65 years of age NVT incidence per 100,000 population increased from a pre vaccine average between 2004/05 and 2005/06 of 16.5 by 48% to 24.4 in 2008/09. In Australia the average NVT incidence in 2002-2004 of 8.4 increased by 75% to 14.6 in 2008. As observed in children, reported levels in the US were generally higher than in the other two countries. In the elderly adjusting these by taking into account only hospitalised cases did not have much impact: even though further ahead in schedule pre vaccination NVT hospitalised IPD in the US only increased from 24.8 cases per 100,000 population to 33.7, a 36% increase. The percentage change in VT (decreasing) and NVT (increasing) adjusted for the time since implementation of the vaccine programme are illustrated in figure 3.13. While there were only slight differences in pre vaccination VT/NVT distribution (56% VT in the US, 68% in Australia and 49% in the UK) and VT reduction was fairly similar the replacement with NVT disease shows a striking difference between the US and the other two countries: For unknown reasons the US did not experience as much serotype replacement.

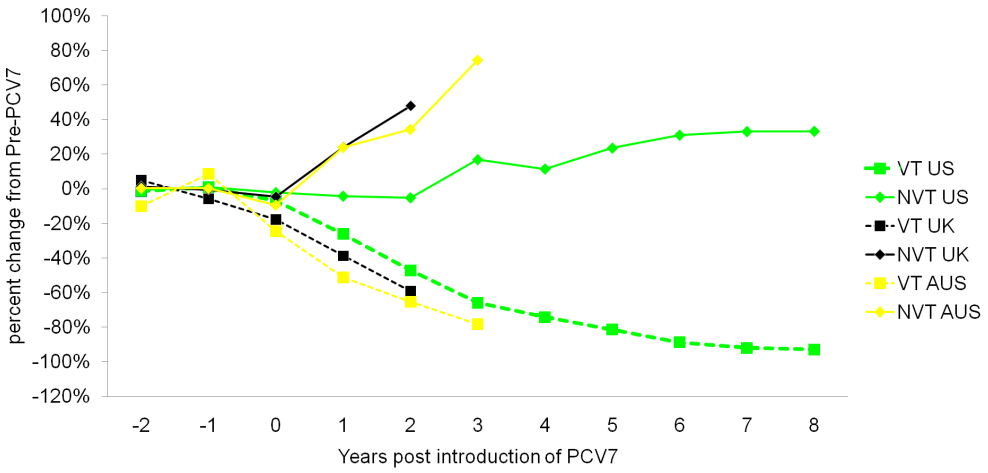


Figure 3.13: Comparison of post-PCV7 experience in the 65+ population. All IPD incidence was standardised to the two year pre vaccination average and the time line was adjusted to represent years since implementation of introduction of PCV7 to the national childhood immunisation program.

Summary

In this chapter the impact of introducing PCV7 into the national immunisation schedule in the United Kingdom is analysed and set in a global perspective. Where pre-vaccination data yield strong trends the interpretation of post vaccination epidemiology is complicated. The analysis presented here shows a possible, however inconclusive, role of potential gradual changes of the sensitivity of pneumococcal surveillance in England and Wales by comparing its pre vaccination epidemiology with these of two control pathogens. Inference from their post vaccination epidemiology shows, that those trends haven't necessarily continued but that making the (implicit) assumption of no trends after the introduction of PCV7 would introduce a bias to a similar extent. This work helped the authors of Miller et al. [17] with their analysis on the impact of PCV7 on IPD.

In Scotland similar trends in the pre-vaccination era were observed. Based on the assumption that this introduces similar caveats to those in England and Wales the analysis of the impact of PCV7 on Scottish IPD is presented stratified into two methodologies: adjusting for trends and assuming unchanged reporting in the post vaccination era. Further, due to previous reports on the volatility of serotype 1 and its major contribution to IPD especially in the adults the analysis was presented both with and without inclusion of serotype 1. One can't really tell which of these ways to analyse the data is most likely to be correct and they should therefore be seen as an estimate of uncertainty introduced by the lack of

understanding of the pneumococcal ecology and epidemiology before introduction of the vaccine.

For both England & Wales and Scotland and in fact, everywhere where it has been introduced, PCV7 has led to a substantial reduction in IPD due to the targeted serotypes in individuals who have been eligible for vaccination. However, discrepancies are reported in the size of the indirect effects (herd immunity and serotype replacement). Section 3.1.2 shows, that independently of the method employed for analysing the data, there is little evidence in Scotland for the, otherwise widely reported, serotype replacement. That is except for the elderly population where serotype replacement balances out the benefits from the indirect protection against VT IPD. This finding in itself is surprising since it is commonly assumed that serotype replacement is driven by the vacation of an ecological niche in the population most responsible for transmission (children) and is passed on from there. However, stochastic effects arising from a small number of cases in Scotland could have contributed to that finding.

Furthermore, the findings about IPD in England and Wales were related to carriage prevalence from pre- and post-vaccination studies conducted in similar populations in England. The analysis showed that, since serotype replacement in carriage was complete, the observed reduction in IPD must have been due to differences in serotype specific invasiveness. The strain-specific estimates for invasiveness were calculated and used to infer that the higher valent vaccines (PCV10/13) include mainly serotypes with high case:carrier ratios which suggests that even in the event of full serotype replacement PCV10/13 are likely to further reduce the burden of invasive pneumococcal disease.

The interpretation of vaccine effects is complicated by the lack of understanding of pre-vaccination epidemiology, which itself is driven by the lack of understanding of the ecology of the pneumococcus. Specifically the mechanisms of competition and coexistence are poorly understood. The subsequent chapters analyse the underlying dynamics of transmission of the pneumococcus by creating and evaluating different modelling approaches and concepts. In particular the next chapter analyses existing model structures and evaluates their feasibility in appropriately reflecting the transmission dynamics of the pneumococcus.

Bibliography

[1] Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. (2003) Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. The New England journal of medicine 348: 1737–46.

[2] Huang SS, Hinrichsen VL, Stevenson AE, Rifas-Shiman SL, Kleinman K, et al. (2009) Continued impact of pneumococcal conjugate vaccine on carriage in young children. Pediatrics 124: e1–11.

[3] Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, et al. (2010) Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. Journal of Infectious Diseases 201: 32–41.

[4] Trotter CL, Waight P, Andrews NJ, Slack M, Efstratiou A, et al. (2010) Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. Journal of Infection 60: 200–208.

[5] Lamb K (2010) Modelling Genetic Effects in the Transmission of Pneumococcal Carriage. Ph.D. thesis, University of Strathclyde.

[6] Flasche S, Slack M, Miller E (2011) Long term trends introduce a potential bias when evaluating the impact of the pneumococcal conjugate vaccination programme in England and Wales. Eurosurveillance 16: 1–6.

[7] Rodenburg GD, de Greeff SC, Jansen AGCS, de Melker HE, Schouls LM, et al. (2010) Effects of pneumococcal conjugate vaccine 2 years after its introduction, the Netherlands. Emerging Infectious Diseases 16: 816–823.

[8] WHO (2010) Changing epidemiology of pneumococcal serotypes after introduction of conjugate vaccine: July 2010 report. Weekly Epidemiological

Record 85: 434–436.

[9] Ihekweazu CA, Dance DAB, Pebody R, George RC, Smith MD, et al. (2008) Trends in incidence of pneumococcal disease before introduction of conjugate vaccine: South West England, 1996-2005. Epidemiology and infection 136: 1096–102.

[10] Health Protection Agency (2010) Guide for Diagnostic Laboratories. Technical report. URL http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1194947381307.

[11] Black S (2010) The Volatile Nature of Pneumococcal Serotype Epidemiology: Potential for Misinterpretation. The Pediatric Infectious Disease Journal 29: 301–303.

[12] Harboe ZB, Benfield TL, Valentiner-Branth P, Hjuler T, Lambertsen L, et al. (2010) Temporal trends in invasive pneumococcal disease and pneumococcal serotypes over 7 decades. Clinical Infectious Diseases 50: 329–337.

[13] HPA (2012). List of notifiable diseases. URL http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/NotificationsOfInfectiousDiseases/ListOfNotifiableDiseases/.

[14] (2009). Childhood immunisation uptake rates, quarter and year ending 31 March 2009. URL http://www.isd.scot.nhs.uk/isd/5954.html.

[15] Jefferies JM, Smith AJ, Edwards GFS, McMenamin J, Mitchell TJ, et al. (2010) Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine. Journal of clinical microbiology 48: 87–96.

[16] Bonferroni CE (1936) Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8: 3–62.

[17] Miller E, Andrews NJ, Waight PA, Slack MP, George RC (2011) Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. The Lancet Infectious Diseases 11: 760–768.

[18] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[19] Brueggemann AB, Griffiths DT, Meats E, Peto T, Crook DW, et al. (2003) Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. Journal of Infectious Diseases 187: 1424–1432.

[20] Weinberger DM, Trzccinski K, Lu YJ, Bogaert D, Brandes A, et al. (2009) Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathogens 5: e1000476.

[21] Black S, Shinefield H, Fireman B, Lewis E, Ray P, et al. (2000) Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. The Pediatric infectious disease journal 19: 187–195.

[22] Dagan R, Melamed R, Muallem M, Piglansky L, Greenberg D, et al. (1996) Reduction of nasopharyngeal carriage of pneumococci during the second year of life by a heptavalent conjugate pneumococcal vaccine. Journal of Infectious Diseases 174: 1271–1278.

[23] Dagan R, Givon-Lavi N, Zamir O, Sikuler-Cohen M, Guy L, et al. (2002) Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. Journal of Infectious Diseases 185: 927–936.

[24] Beutels P, Thiry N, Damme PV (2007) Convincing or confusing? Economic evaluations of childhood pneumococcal conjugate vaccination–a review (2002-2006). Vaccine 25: 1355–1367.

[25] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, et al. (2010) Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC infectious diseases 10: 90.

[26] Wals PD, Robin E, Fortin E, Thibeault R, Ouakki M, et al. (2008) Pneumonia after implementation of the pneumococcal conjugate vaccine program

in the province of Quebec, Canada. The Pediatric infectious disease journal 27: 963–968.

[27] Vestrheim DF, Hø iby EA, Aaberge IS, Caugant DA (2010) Impact of a pneumococcal conjugate vaccination program on carriage among children in Norway. Clinical and Vaccine Immunology 17: 325–334.

[28] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[29] Sleeman KL, Griffiths D, Shackley F, Diggle L, Gupta S, et al. (2006) Capsular serotype-specific attack rates and duration of carriage of *Streptococcus pneumoniae* in a population of children. Journal of Infectious Diseases 194: 682–688.

[30] Weinberger DM, Harboe ZB, Flasche S, Scott JA, Lipsitch M (2011) Prediction of Serotypes Causing Invasive Pneumococcal Disease in Unvaccinated and Vaccinated Populations. Epidemiology 22: 199–207.

[31] Simpson EH (1949) Measurement of Diversity. Nature 163: 688.

[32] Hanage WP, Finkelstein JA, Huang SS, Pelton SI, Stevenson AE, et al. (2010) Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete. Epidemics 2: 80–84.

[33] HPA (2010). Pneumococcal disease. URL `http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Pneumococcal/`.

[34] Salisbury D (2010) Introduction of Prevenar 13 into the Childhood Immunisation Programme. Technical report, Department of Health, London. URL `http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_112192.pdf`.

[35] Turner P, Hinds J, Gould K, Turner C, Jankhot A, et al. (2010) Conventional techniques for detecting nasopharyngeal pneumococcal carriage significantly underestimate the prevalence of multiple serotype carriage. Tel Aviv, p. 27.

[36] Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JAG (2008) The descriptive epidemiology of *Streptococcus pneumoniae* and Haemophilus influenzae nasopharyngeal carriage in children and adults in Kilifi district, Kenya. The

Pediatric infectious disease journal 27: 59–64.

[37] Hill PC, Townend J, Antonio M, Akisanya B, Ebruke C, et al. (2010) Transmission of *Streptococcus pneumoniae* in rural Gambian villages: a longitudinal study. Clinical infectious diseases 50: 1468–1476.

[38] Brugger SD, Frey P, Aebi S, Hinds J, Muehlemann K (2010) Multiple Colonization with *S. pneumoniae* before and after Introduction of the Seven-Valent Conjugated Pneumococcal Polysaccharide Vaccine. PLoS One 5: e11638.

[39] Gray BM, Converse GM, Dillon HC (1980) Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. Journal of Infectious Diseases 142: 923–33.

[40] Weinberger DM, Harboe ZB, Sanders EAM, Ndiritu M, Klugman KP, et al. (2010) Association of Serotype with Risk of Death Due to Pneumococcal Pneumonia: A Meta-Analysis. Clinical infectious diseases 51: 692–9.

[41] Advisory Committee on Immunization Practices (2001) Updated recommendations on the use of pneumococcal conjugate vaccine in a setting of vaccine shortage. Morbidity and Mortality Weekly Report 50: 1140–2.

[42] Centers for Disease Control and Prevention (2008) Invasive pneumococcal disease in children 5 years after conjugate vaccine introduction–eight states, 1998-2005. Morbidity and Mortality Weekly Report 57: 144–148.

[43] Centers for Disease Control and Prevention (CDC) (2011) Invasive pneumococcal disease and 13-valent pneumococcal conjugate vaccine (PCV13) coverage among children aged 59 months—selected U.S. regions, 2010–2011. Morbidity and mortality weekly report 60: 1477–81.

[44] Pelton SI, Loughlin AM, Marchant CD (2004) Seven valent pneumococcal conjugate vaccine immunization in two Boston communities: changes in serotypes and antimicrobial susceptibility among *Streptococcus pneumoniae* isolates. The Pediatric infectious disease journal 23: 1015–22.

[45] Huang SS, Platt R, Rifas-Shiman SL, Pelton SI, Goldmann D, et al. (2005) Post-PCV7 changes in colonizing pneumococcal serotypes in 16 Massachusetts communities, 2001 and 2004. Pediatrics 116: e408—-e413.

[46] Lexau CA, Lynfield R, Danila R, Pilishvili T, Facklam R, et al. (2005) Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. The Journal of the American Medical Association 294: 2043–2051.

[47] Hsu HE, Shutt KA, Moore MR, Beall BW, Bennett NM, et al. (2009) Effect of pneumococcal conjugate vaccine on pneumococcal meningitis. New England Journal of Medicine 360: 244–256.

[48] Steenhoff AP, Shah SS, Ratner AJ, Patil SM, McGowan KL (2006) Emergence of vaccine-related pneumococcal serotypes as a cause of bacteremia. Clinical infectious diseases 42: 907–914.

[49] Eskola J, Kilpi T, Palmu A, Jokinen J, Haapakoski J, et al. (2001) Efficacy of a pneumococcal conjugate vaccine against acute otitis media. The New England journal of medicine 344: 403–9.

[50] Centre for Disease Control and Prevention (2009) Pneumonia hospitalizations among young children before and after introduction of pneumococcal conjugate vaccine–United States, 1997-2006. Morbidity and Mortality Weekly Report 58: 1–4.

[51] Tsai CJ, Griffin MR, Nuorti JP, Grijalva CG (2008) Changing epidemiology of pneumococcal meningitis after the introduction of pneumococcal conjugate vaccine in the United States. Clinical infectious diseases 46: 1664–1672.

[52] Gomes HDC, Muscat M, Monnet DL, Giesecke J, Lopalco PL (2009) Use of seven-valent pneumococcal conjugate vaccine (PCV7) in Europe, 2001-2007. Eurosurveillance 14: 1–6.

[53] Moszynski P (2011) GAVI rolls out vaccines against child killers to more countries. British Medical Journal 343: d6217–d6217.

[54] Center KJ (2007) Prevenar vaccination: review of the global data, 2006. Vaccine 25: 3085–3089.

[55] Goldblatt D, Southern J, Ashton L, Andrews N, Woodgate S, et al. (2010) Immunogenicity of a reduced schedule of pneumococcal conjugate vaccine in healthy infants and correlates of protection for serotype 6B in the United Kingdom. The Pediatric Infectious Disease Journal 29: 401–5.

[56] Vestrheim DF, Lø volll O, Aaberge IS, Caugant DA, Hø iby EA, et al. (2008) Effectiveness of a 2+1 dose schedule pneumococcal conjugate vaccination programme on invasive pneumococcal disease among children in Norway. Vaccine 26: 3277–3281.

[57] Williams SR, Mernagh PJ, Lee MHT, Tan JT (2011) Changing epidemiology of invasive pneumococcal disease in Australian children after introduction of a 7-valent pneumococcal conjugate vaccine. The Medical journal of Australia 194: 116–20.

[58] Hausdorff WP, Bryant J, Kloek C, Paradiso PR, Siber GR (2000) The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use, part II. Clinical infectious diseases 30: 122–140.

[59] Hausdorff WP, Bryant J, Paradiso PR, Siber GR (2000) Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. Clinical infectious diseases 30: 100–121.

[60] Miller L (2010) Post-PCV serotype changes and variability in observations. In: 7th International Symposium on Pneumococci and Pneumococcal Diseases. Tel Aviv.

[61] WHO Experts (2011). Immunization, Vaccines and Biologicals - SAGE meeting of 8 - 10 November 2011. URL `http://www.who.int/immunization/sage/previous_november2011/en/index.html`.

[62] Bettinger J, Scheifele D, Kellner J (2010) The effect of routine vaccination on invasive pneumococcal infections in Canadian children, Immunization Monitoring Program, Active 2000-2007. Vaccine .

[63] Techasaensiri C, Messina AF, Katz K, Ahmad N, Huang R, et al. (2010) Epidemiology and evolution of invasive pneumococcal disease caused by multidrug resistant serotypes of 19A in the 8 years after implementation of pneumococcal conjugate vaccine immunization in Dallas, Texas. The Pediatric infectious disease journal 29: 294–300.

4 Transmission patterns of the pneumococcus

Outline

In the previous chapter the epidemiology of *S. pneumoniae* in Britain was studied. Increasing trends in disease incidence prior to the introduction of the 7-valent conjugate vaccine into the routine childhood immunisation scheme, and the volatility of serotype 1 contributed uncertainty to the assessment of the vaccine induced changes in the epidemiology of IPD. However, there is good evidence that PCV7 did substantially reduce the burden of VT IPD in vaccinated individuals and through herd immunity in the unvaccinated population as well. Furthermore, the analysis of nasopharyngeal colonisation data pre and post vaccination in England and Wales yielded full serotype replacement in carriage. This highlights the importance of type specific invasiveness and provides evidence that both PCV10 and PCV13 are likely to have further beneficial impact on reducing the burden of invasive pneumococcal disease; even in the event of complete serotype replacement. The analysis showed the need for a better understanding of the ecology of the pneumococcus in order to appropriately predict the impact of higher valent vaccine formulations on the precursor of disease; namely pneumococcal carriage.

In order to build a model of the spread of the pneumococcus, and its prevention through vaccination or other means, the essential mechanisms of transmission and persistence of the pneumococcus in the population need to be identified and parameterised. In this section information on the transmission patterns of the pneumococcus is reviewed both for parameter values and, more importantly, on a structural level. The scientific literature is reviewed for information on the events of serotype- and age-specific acquisition and clearance. The focus here is

the antagonism of competition of the different serotypes and the coexistence of them. Existing models incorporating the mechanisms of competition are identified, reviewed and investigated for steady states of stable coexistence and the explicit and implicit assumptions which grant them.

4.1 Acquisition and clearance

Information on pneumococcal transmission patterns is sparse since epidemiological estimates require longitudinal studies of carriage usually done by nasopharyngeal swabbing which is an unpleasant procedure for the study participant. However an increasing number of these studies have been carried out and provide valuable information on the patterns of acquisition and clearance of *Streptococcus pneumoniae*.

The clearance of the pneumococcus is found to be both age and type specific. Independent of the capsular serotype the duration of carriage (the reciprocal of the clearance rate) is highest in children under one year of age and declines gradually thereafter [1, 2]. The average duration of carriage in a UK setting was found to be 72 days for 0-1 year olds and 28, 18, 17 days for 2-4, 5-17, 18+ year olds respectively [3]. However, there is substantial variation amongst the serotypes. In a study of more than 300 children under two years of age in Oxfordshire serotype specific duration of carriage varied between about 5 weeks for serotypes 7F and 12F and more than 15 weeks for serotypes 6B and 23F which is in line with further studies in England and Wales [4]. Also some serotypes, e.g. 1 and 5, are hardly found in any carriage study conducted in a developed country (compare section 3.1.3) which could indicate an even lower duration of carriage for those. The reasons for differences in clearance rates, and therefore likely differences in carriage prevalence, are various and could include differences in the degree of encapsulation. This provides an advantage in the susceptibility against killing by neutrophils and therefore is correlated with the prevalence of colonisation [5].

Similar to clearance, acquisition of carriage of the pneumococcus is age and serotype specific. Children were found to be about 3-4 times more likely to acquire pneumococcal infections compared to adults [1] and acquisition rates (carriage incidence) differ greatly depending on the capsule structure [4, 6].

Vaccination with the seven valent conjugate vaccine was proven effective in reducing the prevalence of pneumococcal colonisation of vaccine-types which confers additional benefits through herd immunity. This could be achieved in at least two different ways, either through reducing acquisition rates or through increasing clearance rates. However, it has been shown that vaccination only has a significant effect on reducing pneumococcal acquisition [7, 8]. Since invasive disease is believed to follow shortly after acquisition [9] an increased clearance rate would have been unlikely to protect directly against invasive disease, however, it could still have conferred indirect protection through reducing transmission.

While for invasive pneumococcal disease there is a pronounced peak every year around winter time, which is thought to be partly the results of viral infections (see section 2.7) increasing the susceptibility of the immune system against invasion of the pneumococcus [10], mixed evidence on the seasonality of pneumococcal acquisition exists. No seasonality in carriage prevalence in regions with pronounced winter peaks in IPD was detected is most studies [11–13], however Lakshman and colleagues [14] reported around 40% increased carriage prevalence in winter and Gray and colleagues [11] reported significant seasonal patterns in the acquisition of the pneumococcus which were outbalanced by seasonally affected clearance rates and therefore did not lead to seasonally varying carriage rates. Under the assumption that clearance rates do not significantly affect progression to disease the seasonality in IPD could still come from seasonality in acquisition rates even though they are balanced out by seasonal clearance rates in carriage prevalence.

4.2 Replacement and competition

The introduction of the 7-valent pneumococcal conjugate vaccine has led to a great reduction in invasive disease caused by the serotypes included in the vaccine formulation in vaccinated individuals and even in unvaccinated age groups through herd immunity effects. However this benefit has been partly offset by increasing incidence of NVT IPD. As discussed in section 3.2, the extent of the replacement varies over different settings. The issue of replacement is even more definite in pneumococcal carriage. In various settings almost complete replacement of VT with NVT carriage was reported after introduction of PCV7 to the childhood immunisation scheme [15–18]. Assessment biases which could alter the

observed level of serotype replacement include: changes in reporting behaviour, demographic changes, natural fluctuations of the prevalence of certain serotypes, insufficient detection methods and antibiotic susceptibility. These are discussed in a review by Weinberger and colleagues [19] who conclude, that they might introduce a bias but are unlikely to be sufficient to dismiss the causal relationship between vaccine introduction and increase in NVT carriage and disease.

The observation of vaccine induced replacement in the pneumococcal population indicates competition amongst pneumococcal serotypes. By vaccinating against the most prevalent serotypes in pneumococcal carriage the NVTs face less competition and therefore invade the vacated niche. Yet, the mechanisms of competition are not fully understood. Auranen and colleagues explored whether competition reduces the acquisition of new serotypes or increases the clearances of the existing ones in a longitudinal study in day-care children [20]. They found evidence for the most common serotypes, that colonised children have about a 90% reduced rate of acquisition compared to non carrying children. The rate of clearance however, was not significantly different.

Different approaches exist to model competition amongst serotypes of the pneumococcus and infer changes due to future vaccine formulations. A rather elementary method infers the changes in IPD by assuming that replacement in carriage is complete, that the NVT prevalence increases proportionately to pre vaccination levels and that invasiveness remains constant [21]. This approach was also used in section 3.1.3. In the following approaches using dynamic models to infer the impact of vaccination in the presence of competition will be evaluated.

4.3 Modelling approaches

It was established in chapter 3 that key aspects of pneumococcal ecology following conjugate vaccination are herd immunity and serotype replacement which is believed to arise from the loss of competition, induced by the targeted serotypes, on the other types. However, the actual mechanisms leading to competition between strains remain unknown. Different approaches approaches have studied the general interaction of multiple pathogens and how their ecology could impact on vaccination [22–24] . To reflect in particular the indirect effects observed after conjugate vaccination against *S. pneumonia* various approaches in non dynamic models exist. Some just disregard them [1, 25] others assume that indirect

effects following introduction of PCV13 are similar to those observed following PCV7 [26]. However, non-dynamic approaches cannot explicitly model the dynamics leading to indirect effects and, as a consequence cannot capture likely differences of observed and unobserved dynamics, those arising from differences in carriage prevalence as lower prevalence of vaccine serotypes is likely to cause less serotype replacement. Therefore, non-dynamic modelling approaches as well as approaches considering the pneumococcus as a whole (no competing groups of serotypes or resistant strains; e.g. [27, 28]) are disregarded in the remainder of this chapter. Dynamic modelling approaches, which include competition of two pneumococcal serotype groups, can be largely reduced to three approaches. The corresponding model diagrams, equations and stability analysis are provided in the respective sections 4.3.2, 4.3.3 and 4.3.4.

Triangle model without strain interaction. This model is an SIS model consisting of three different compartments: the infected (and infectious) with either one of the serotypes and the susceptible compartment. It represents the simplest form of competition: competitive exclusion, that is two strains compete for colonisation of a host where for the time of the infection the host becomes fully resistant against infection with the other type (heterologous protection) and against re-infection with the same type (homologous protection). Amongst others this model type was mostly used to study the behaviour of antibiotic susceptibility of the pneumococcus [29–35]. With improving detection techniques (see section 2.5.2) a commonly observed feature of pneumococci is the co-colonisation of the host which provides evidence for imperfect protection against additional colonisation through colonisation and which cannot be reflected in this model structure.

Triangle model with strain interaction. The inclusion of strain interaction extends the previous triangle model by adding the possibility to go directly from the state of infection with one strain to infection with the other (with a reduced probability of infection). This event can be interpreted as the serotypes outcompeting each other for space in the nasopharynx with only one strain being able to grow at a time. As with the model above this model ignores multiple colonisation as a relevant factor for the transmission dynamics. However, the assumption that a new colonisation with serotype two given infection with serotype one results in a reduced chance for strain one to invade and, more importantly, to be transmitted makes the inclusion of a compartment for dual colonisation unnecessary. This model was used to

estimate serotype-specific transmission parameters, such as duration of carriage and acquisition rates, from a longitudinal carriage study in the UK [4] and to model multi type transmission of pneumococci in families [36].

Diamond-type model. The diamond-type model introduces the possibility of being colonised with two strains at a time. The competition between strains is reflected by a reduced chance of additional colonisation while colonised with the heterologous strain and by denying additional colonisation with the homologous strain. Its theory was introduced in 1997 [37] and since then used for models of penicillin resistance [38], predictions of vaccination impacts [3, 39, 40] and estimation of competition between serotypes from longitudinal carriage studies [20]. However, concerns have been raised about the implicit assumptions that lead to coexistence in this framework [41] and models have been proposed which, for example, allowed the co-colonisation with a homologous type.

In the following these models will be analysed to see their specific effects of competition on steady state solutions of stable coexistence. But first the theoretical concept of a neutral model will be introduced.

4.3.1 Neutral models

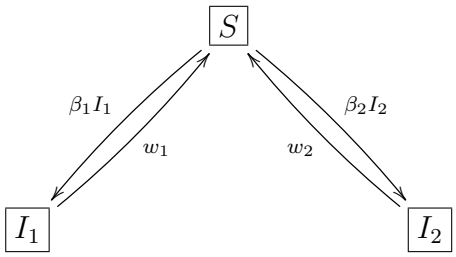
There are more than 90 different serotypes of pneumococcus identified. A huge variety of them are circulating in a population at the same time; most studies of nasopharyngeal colonisation report more than 30 different serotypes isolated in the study population [15, 16, 18]. Even though the prevalence of pneumococcal serotypes is remarkably different (which probably infers different levels of fitness) they coexist. When employing a model that allows for competition of strains and inferring coexistence the mechanisms granting coexistence should be explicit rather than a hidden dynamic which is promoting coexistence arbitrarily. Lipsitch and colleagues therefore defined the framework of a neutral model [41]. When applied to indistinguishable strains a neutral null model should fulfil two criteria:

ecological neutrality The dynamics of the ecological variables, the number of uninfected hosts (N_0) and the number of people infected with 1,2,... strains ($N_1, N_2, \ldots$), should depend on the ecological state variables but, given these, should be independent of the identity of the particular strains involved. Note that the ecological variables sum up to the total population

size (i.e. $N = \sum_i N_i$)

population genetic neutrality There should not be a stable equilibrium frequency of the strains in the model; rather, it should be possible to choose initial conditions to guarantee an arbitrary frequency of strains that remains constant for all time t ≥ 0.

4.3.2 Triangle model without strain interaction



Consider a basic triangle model with susceptible (S), infected with strain 1 (I_1) and infected with strain 2 (I_2) and the overall population size $1 = N = S + I_1 + I_2$ (setting $N = 1$ can be interpreted as transforming the number of susceptibles and infective to proportions respective to the total population size):

$$\begin{aligned} S' &= -\beta_1 S I_1 - \beta_2 S I_2 + w_1 I_1 + w_2 I_2 \\ I_1' &= +\beta_1 S I_1 - w_1 I_1 \\ I_2' &= +\beta_2 S I_2 - w_2 I_2 \end{aligned} \tag{4.3.1}$$

with β_1 and β_2 representing the population contact/transmission patterns and w_1 and w_2 the strain specific rates of waning infectiousness and protection against new acquisition. As with most biological models only parameters exceeding zero are feasible here. The strain specific reproduction number is $R_{0i} = \frac{\beta_i}{w_i},\ i \in \{1, 2\}$ (see section 1.2). For indistinguishable strains ($\beta = \beta_1 = \beta_2$, $w = w_1 = w_2$) this model fulfils the first criteria of a neutral model (ecological neutrality) following Lipsitch et al. [41]:

$$\begin{aligned} N_0' &= -\beta N_0 N_1 + w N_1 \\ N_1' &= +\beta N_0 N_1 - w N_1 \end{aligned}$$

with $N_0 = S$ and $N_1 = I_1 + I_2$. The steady states (where the system is at equilibrium: $S' = I_1' = I_2' = 0$) result from:

$$\begin{aligned} 0 &= -\beta_1 S I_1 - \beta_2 S I_2 + w_1 I_1 + w_2 I_2 \\ 0 &= +\beta_1 S I_1 - w_1 I_1 \\ 0 &= +\beta_2 S I_2 - w_2 I_2 \end{aligned}$$

Considering the equation for I_1:

$$\begin{aligned} 0 &= (\beta_1 S - w_1) I_1 \\ \Rightarrow 0 &= I_1 \text{ or } 0 = (\beta_1 S - w_1) \end{aligned}$$

The analogue holds for I_2 leading to:

$$\begin{aligned} I_1 = 0, I_2 = 0 &\Rightarrow S = 1 \\ I_1 = 0, S = w_2/\beta_2 &\Rightarrow I_2 = 1 - w_2/\beta_2 \\ S = w_1/\beta_1, I_2 = 0 &\Rightarrow I_1 = 1 - w_1/\beta_1 \\ S = w_1/\beta_1, S = w_2/\beta_2 &\text{ ,only holds when: } \frac{w_1}{\beta_1} = \frac{w_2}{\beta_2} \end{aligned}$$

So the rates of susceptibles and infected with either strains at steady state are:

$$(S^s, I_1^s, I_2^s) \in \left\{ (1,0,0), \left(\frac{1}{R_{02}}, 0, 1 - \frac{1}{R_{02}} \right), \left(\frac{1}{R_{01}}, 1 - \frac{1}{R_{01}}, 0 \right) \right\}.$$

Note that since this is a population model which is only feasible for $S, I_1, I_2 \geq 0$ the non disease free equilibria only exist if the respective reproduction number is bigger than 1. An additional steady state exists if $\frac{\beta_1}{w_1} = \frac{\beta_2}{w_2}$ (the reproduction number is the same for both strains):

$$(S^s, I_1^s, I_2^s) = \left(\frac{1}{R_{01}}, I_1^s, 1 - I_1^s - \frac{1}{R_{01}} \right) \text{ ,} \forall I_1^s : 0 < I_1^s < 1 - \frac{1}{R_{01}}.$$

Hence only in case $R_{01} = R_{02}$ there exists a set of steady states allowing different levels of prevalence of both strains for the equilibrium. Therefore the second criterion of a neutral model (the population genetic neutrality) is fulfilled and the model is neutral. This shows that stable coexistence of two strains in this model is limited to the case where $R_{01} = R_{02}$ but with I_1^s not necessarily equal to I_2^s. In the rest of the parameter space there is no equilibrium with coexistence of both types. Further the necessity of $S^s, I_1^s, I_2^s \geq 0$ reveals the need for $R_{01} > 1$

or $R_{02} > 1$ (compare the threshold theorem in section 1.2) for the existence of a single strain steady state and $R_{01} = R_{02} > 1$ for coexistence of both strains.

Stability analysis

Local stability (Lyapunov stability) of a steady state is ensured if the Jacobian matrix of the differential equations at steady state has only eigenvalues with non positive real parts. This ensures that if the solution of the system of differential equations is in a neighbourhood around an equilibrium point it won't diverge from it. Further, asymptotic stability is granted by negative real parts of the eigenvalues and ensures convergence of the system to the equilibrium point once it is close enough. The Jacobian matrix corresponding to equation 4.3.1 is:

$$\begin{pmatrix} -\beta_1 I_1 - \beta_2 I_2 & -\beta_1 S + w_1 & -\beta_2 S + w_2 \\ \beta_1 I_1 & \beta_1 S - w_1 & 0 \\ \beta_2 I_2 & 0 & \beta_2 S - w_2 \end{pmatrix}.$$

Disease free equilibrium: For the steady state $B_1 = (1, 0, 0)$ the Jacobian matrix is:

$$\begin{pmatrix} 0 & -\beta_1 + w_1 & -\beta_2 + w_2 \\ 0 & \beta_1 - w_1 & 0 \\ 0 & 0 & \beta_2 - w_2 \end{pmatrix}.$$

The eigenvalues can be read off the diagonal: $EV_1 = 0$, $EV_2 = \beta_1 - w_1$, $EV_3 = \beta_2 - w_2$. For local stability in B_1 therefore $R_{01} \leq 1$ & $R_{02} \leq 1$ have to hold and ensure that the disease free equilibrium will not give way to invasion by either strain.

Dominance of strain 1: For the steady state $B_2 = \left(\frac{1}{R_{01}}, 1 - \frac{1}{R_{01}}, 0\right)$ the Jacobian matrix is:

$$\begin{pmatrix} -\beta_1 + w_1 & 0 & -\beta_2 \frac{w_1}{\beta_1} + w_2 \\ \beta_1 - w_1 & 0 & 0 \\ 0 & 0 & \beta_2 \frac{w_1}{\beta_1} - w_2 \end{pmatrix}.$$

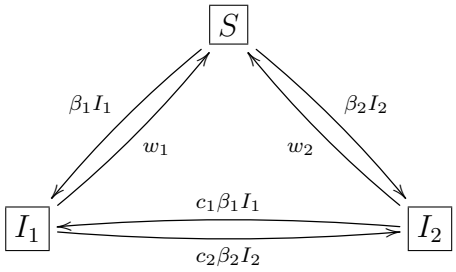
Making use of the Laplace extension the eigenvalues can be read of the diagonal: $EV_1 = 0$, $EV_2 = -\beta_1 + w_1$, $EV_3 = \beta_2 \frac{w_1}{\beta_1} - w_1$. This shows the need for $R_{01} \geq 1$ and $R_{02} \leq R_{01}$ for B_2 to be locally stable; in other words: if the reproduction

number of the first strain is not smaller than one and the reproduction number of the second strain is smaller than that, stable existence of the first strain in the population is ensured and invasion of the second strain is inhibited.

Dominance of strain 2: Stability of $B_3 = \left(\frac{1}{R_{02}}, 0, 1 - \frac{1}{R_{02}}\right)$ can be inferred easily because the model 4.3.1 is symmetric: B_3 is locally stable in case $R_{02} \geq 1$ and $R_{01} \leq R_{02}$.

Coexistence: If B_4 does not arise from B_2 and B_3 through transcritical bifurcation, one can infer the domain for stability of the steady state of coexistence [42]. Assuming that this is the case, B_4 exists and is stable on the domain, where the steady states of a single type's survival become unstable (or at least no longer asymptotically stable) through invasion by the other type. This suggests the existence and stability of a steady state of coexistence if $R_{01} = R_{02} \geq 1$, which complies with what was calculated earlier ($B_4 = \left(\frac{1}{R_{01}}, I_1^s, 1 - I_1^s - \frac{1}{R_{01}}\right)$,if $R_{01} = R_{02}$) but further suggests stability.

4.3.3 Triangle model with strain interaction



Consider an extended version of the basic triangle model now where an infected individual can clear his infection by acquiring the other strain (for another interpretation see the beginning of the section 4.3):

$$S' = -\beta_1 S I_1 - \beta_2 S I_2 + w_1 I_1 + w_2 I_2 \tag{4.3.2}$$

$$I_1' = +\beta_1 S I_1 - w_1 I_1 - c_2\beta_2 I_1 I_2 + c_1\beta_1 I_2 I_1 \tag{4.3.3}$$

$$I_2' = +\beta_2 S I_2 - w_2 I_2 + c_2\beta_2 I_1 I_2 - c_1\beta_1 I_2 I_1 \tag{4.3.4}$$

with the same parameters as before and $c_{1,2} \in [0, 1]$ being the competition parameter. Note that for competitive exclusion $c_1 = c_2 = 0$ this model is the same

as the previous one. Like for the simple triangle model the strain specific reproduction number is $R_{0i} = \frac{\beta_i}{w_i},\ i \in \{1, 2\}$. For indistinguishable strains this model fulfils the first criteria of a neutral model (ecological neutrality) following Lipsitch et al. [41]:

$$\begin{aligned} N_0' &= -\beta N_0 N_1 + w N_1 \\ N_1' &= +\beta N_0 N_1 - w N_1 + 0 \end{aligned}$$

For $I_1 \vee I_2 \equiv 0$ the model reduces to the triangle model discussed in the first paragraph of section 4.3.2 so the steady states are derived analogously:

$$\begin{aligned} I_1 = 0, I_2 = 0 &\Rightarrow S = 1 \\ I_1 = 0, I_2 \neq 0 &\Rightarrow S = w_2/\beta_2, I_2 = 1 - w_2/\beta_2 \\ I_1 \neq 0, I_2 = 0 &\Rightarrow S = w_1/\beta_1, I_1 = 1 - w_1/\beta_1 \end{aligned}$$

Only the case $I_1 \wedge I_2 \neq 0$ needs some further investigation: In the case when $c_1\beta_1 - c_2\beta_2$ is zero this model limits to the model investigated previously so a steady state with coexistence is only found for $R_{01} = R_{02} > 1$. In particular, this is the case for indistinguishable strains; therefore the triangle model with strain interaction also fulfils the criteria of population genetic neutrality and is neutral. However, for a complete steady state assessment the case when $c_1\beta_1 - c_2\beta_2 \neq 0$ needs investigation. Substitute $\psi = c_1\beta_1 - c_2\beta_2$. Then

$$\begin{aligned} 4.3.3 \Rightarrow 0 &= \beta_1 S I_1 - w_1 I_1 + \psi I_2 I_1 \\ &\overset{I_1 \neq 0}{=} \beta_1 S - w_1 + \psi I_2 \\ \Rightarrow I_2 &= -\frac{\beta_1 S - w_1}{\psi} = -\frac{(R_{01} S - 1) w_1}{\psi} \qquad (4.3.5) \\ 4.3.4 \Rightarrow 0 &= \beta_2 S I_2 - w_2 I_2 - \psi I_1 I_2 \\ &\overset{I_2 \neq 0}{=} \beta_2 S - w_2 - \psi I_1 \\ \Rightarrow I_1 &= \frac{\beta_2 S - w_2}{\psi} = \frac{(R_{02} S - 1) w_2}{\psi}. \qquad (4.3.6) \end{aligned}$$

With $1 = N = S + I_1 + I_2$ the equilibrium completes to:

$$\begin{aligned} S =& 1 - I_1 - I_2 \\ =& 1 - \frac{-\beta_1 S + w_1 + \beta_2 S - w_2}{\psi} \\ (\psi - \beta_1 + \beta_2) S =& \psi - w_1 + w_2 \\ S =& \frac{\psi - w_1 + w_2}{\psi - \beta_1 + \beta_2} = \frac{\psi - w_1 + w_2}{\psi - R_{01} w_1 + R_{02} w_2} \end{aligned} \tag{4.3.7}$$

So the steady states for $\psi \neq 0$ are:

$$\begin{aligned} (S^s, I_1^s, I_2^s) \in \Bigg\{ & (1,0,0), \left(\frac{1}{R_{02}}, 0, 1 - \frac{1}{R_{02}} \right), \left(\frac{1}{R_{01}}, 1 - \frac{1}{R_{01}}, 0 \right), \\ & \left(\frac{\psi - w_1 + w_2}{\psi - R_{01} w_1 + R_{02} w_2}, \frac{(R_{02} S - 1) w_2}{\psi}, -\frac{(R_{01} S - 1) w_1}{\psi} \right) \Bigg\} \end{aligned}$$

As before the single type survival is limited to $R_{01,02} > 1$ to fulfil the need for positivity of the solutions. The conditions for the state of coexistence are less obvious. They include $c_1\beta_1 - c_2\beta_2 \neq 0$ and $c_1\beta_1 - c_2\beta_2 - \beta_1 + \beta_2 \neq 0$. As mentioned before the case $c_1\beta_1 - c_2\beta_2 = 0$ reduces the model to the normal triangle one discussed in detail in the section 4.3.2 and therefore needs no further investigation. The restriction $c_1\beta_1 - c_2\beta_2 - \beta_1 + \beta_2 \neq 0$ can be transformed to

$$\begin{aligned} c_1\beta_1 - c_2\beta_2 - \beta_1 + \beta_2 = (-1 + c_1)\beta_1 + (1 - c_2)\beta_2 \neq 0 \\ \Rightarrow \quad (1 - c_1)\beta_1 \neq (1 - c_2)\beta_2. \end{aligned}$$

Violation of this criteria again includes indistinguishable strains but also denies coexistence to a wider set of parameter choices.

Stability analysis

The Jacobian matrix is:

$$\begin{pmatrix} -\beta_1 I_1 - \beta_2 I_2 & -\beta_1 S + w_1 & -\beta_2 S + w_2 \\ \beta_1 I_1 & \beta_1 S - w_1 - c_2\beta_2 I_2 + c_1\beta_1 I_2 & -c_2\beta_2 I_1 + c_1\beta_1 I_1 \\ \beta_2 I_2 & c_2\beta_2 I_2 - c_1\beta_1 I_2 & \beta_2 S - w_2 + c_2\beta_2 I_1 - c_1\beta_1 I_1 \end{pmatrix}.$$

Disease free equilibrium: In $B_1 = (1, 0, 0)$ the Jacobian matrix is:

$$\begin{pmatrix} 0 & -\beta_1 + w_1 & -\beta_2 + w_2 \\ 0 & \beta_1 - w_1 & 0 \\ 0 & 0 & \beta_2 - w_2 \end{pmatrix}.$$

Therefore the eigenvalues are $EV_1 = 0$, $EV_2 = \beta_1 - w_1$, $EV_3 = \beta_2 - w_2$ and impose local stability for B_1 in case $R_{01} \leq 1$ and $R_{02} \leq 1$.

Dominance of strain 1: In $B_2 = \left(\frac{1}{R_{01}}, 1 - \frac{1}{R_{01}}, 0\right)$ the Jacobian matrix is:

$$\begin{pmatrix} -\beta_1 + w_1 & 0 & -\beta_2 \frac{w_1}{\beta_1} + w_2 \\ \beta_1 - w_1 & 0 & c_2\beta_2 + c_2\beta_2 \frac{w_1}{\beta_1} + c_1\beta_1 - c_1 w_1 \\ 0 & 0 & \beta_2 \frac{w_1}{\beta_1} - w_2 + c_2\beta_2 - c_2\beta_2 \frac{w_1}{\beta_1} - c_1\beta_1 + c_1 w_1 \end{pmatrix}.$$

Bearing in mind the Laplace expansion one can read off the eigenvalues again: $EV_1 = 0, \quad EV_2 = -\beta_1 + w_1, \quad EV_3 = \beta_2 \frac{w_1}{\beta_1} - w_2 + c_2\beta_2 - c_2\beta_2 \frac{w_1}{\beta_1} - c_1\beta_1 + c_1 w_1$. Therefore B_2 is locally stable in case

$$R_{01} \geq 1, \text{ and} \tag{4.3.8}$$

$$R_{02} \leq R_{01} \frac{w_2 + c_1 w_1 (R_{01} - 1)}{w_2 + c_2 w_2 (R_{01} - 1))}. \tag{4.3.9}$$

Similar to section 4.3.2 here the restriction in equation 4.3.8 ensures the persistence of strain one in the population and the restriction in equation 4.3.9 prevents the invasion of the other strain. Note that if one allows for variable levels of competition instead of assuming competitive exclusion as in section 4.3.2 the restriction for stability of the steady state preventing invasion of the second strain becomes no longer dependent only on the other types reproduction number but includes the degree of competition as well.

Dominance of strain 2: Similar to what was shown before one finds that

$$R_{02} \geq 1, \text{ and}$$

$$R_{01} \leq R_{02} \frac{w_1 + c_2 w_2 (R_{02} - 1)}{w_1 + c_1 w_1 (R_{02} - 1))}.$$

need to hold to ensure local stability of $B_3 = \left(\frac{1}{R_{02}}, 0, 1 - \frac{1}{R_{02}}\right)$.

Coexistence: As in section 4.3.2 it is assumed that in the set where B_2 and B_3 are not stable and prone to invasion of the respective other strain the steady state of coexistence exists and is stable (again, this assumption may not be true for some cases, see Zhang et al. [43]). Hence, this model provides stable coexistence of both strains in B_4 on the domain C described by:

$$
\begin{aligned}
R_{01} &\geq 1 \\
R_{01} &\geq R_{02}\frac{w_1 + c_2 w_2(R_{02} - 1)}{w_1 + c_1 w_1(R_{02} - 1))} \\
R_{02} &\geq 1 \\
R_{02} &\geq R_{01}\frac{w_2 + c_1 w_1(R_{01} - 1)}{w_2 + c_2 w_2(R_{01} - 1))}
\end{aligned}
$$

This includes the special case $c_1 w_1 = c_2 w_2$ (and the obvious case of no competition $c_{1,2} = 0$) which reduces the constraint to $R_{01} = R_{02}$ for which the existence of a steady state was shown earlier.

To test the existence of steady states of coexistence numerically the triangle model with strain interaction was implemented and numerically solved using the explicit Euler method with a time-step of 1. For a fixed set of parameters 1000 model simulations were run. Each simulation sampled R_1 and R_2 from a uniform distribution between 0 and 5 and was run until the sum of the Euclidean differences of S, I_1, I_2 at times t and $t-1$ fell under an error threshold of $1 * 10^{-7}$; the model was then considered to be in a steady state. A strain was considered extinct if at steady state less than 5‰of the population were infected (other thresholds yielded similar results). The simulations confirmed the hypothesis of the derived domain granting a stable steady state of coexisting strains for all tested parameter sets (see Figure 4.1). Furthermore, by its design the implemented model suggests global stability of the triangle model on C as well. Although analytically laborious it is easy to show numerically that the conditions derived for stable coexistence on C correspond to the requirement of positivity of the steady state granting coexistence. The parameter space of coexistence is generally narrow. Only $c_{1,2} \gg w_{1,2}$ together with a substantial difference in the competition induced by either strains increases the parameter space for stable coexistence.

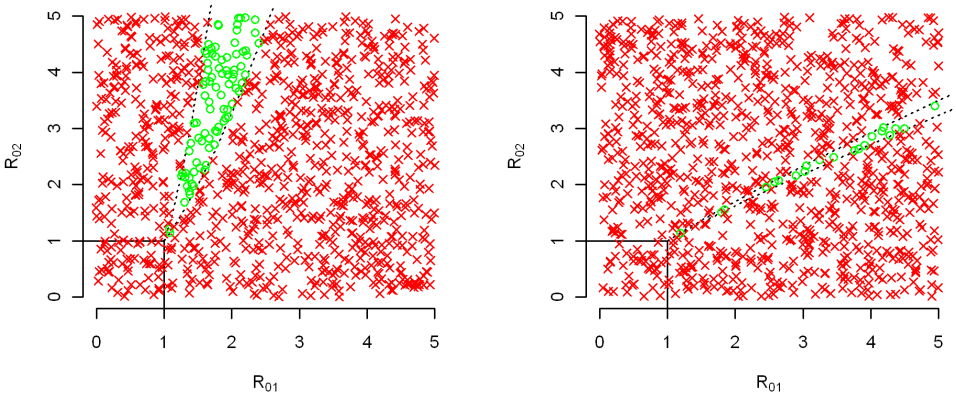
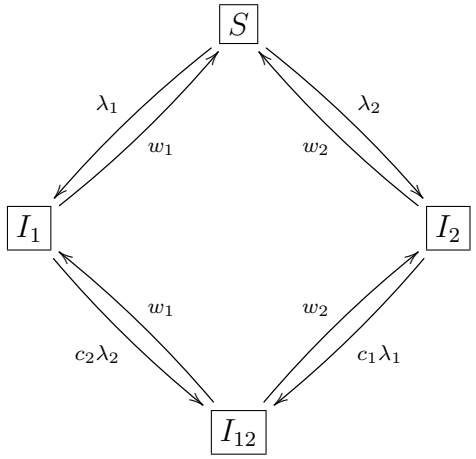


Figure 4.1: Numerical model results for the test of stable coexistence. The dashed line marks the domain C. Model simulations which resulted in coexistence of both strains are represented by green circles and the extinction of at least one strain by red crosses. The figure consists of 1000 simulations with uniformly distributed R_{01} and R_{02}, the other parameter values are: $w_1 = 1/60, w_2 = 1/60, c_1 = 0.8, c_2 = 0.1$ (left), $w_1 = 1/6, w_2 = 1/6, c_1 = 0.3, c_2 = 0.6$ (right) and the initial values: $(S, I_1, I_2) = (0.2, 0.4, 0.4)$.

4.3.4 Diamond-type model



Now consider a diamond type model incorporating a state of co-colonisation as proposed by Lipsitch [37]. Denote the susceptible by S, the infected with strain 1 by I_1, the infected with strain 2 by I_2 and carriers of both strains with by I_{12}.

The equations are:

$$\begin{aligned} S' &= -(\lambda_1 + \lambda_2)S + w_1 I_1 + w_2 I_2 + \mu - \mu S \\ I_1' &= \lambda_1 S - w_1 I_1 - c_2 \lambda_2 I_1 + w_2 I_{12} - \mu I_1 \\ I_2' &= \lambda_2 S - w_2 I_2 - c_1 \lambda_1 I_2 + w_1 I_{12} - \mu I_2 \\ I_{12}' &= c_2 \lambda_2 I_1 + c_1 \lambda_1 I_2 - (w_1 + w_2) I_{12} - \mu I_{12} \end{aligned} \tag{4.3.10}$$

with $\lambda_1 = \beta_1(I_1 + qI_{12})$ and $\lambda_2 = \beta_2(I_2 + qI_{12})$ using the same name conventions for parameters as before. Here the parameter $q \in [0, 1]$ scales the infectivity of the dually colonised; $q = \frac{1}{2}$ corresponds to the dual infected being as infective (in total) as the individuals infected with a single strain. This is assuming that a person dually colonised has equal chances of transmitting either of the strains which otherwise could be overcome by splitting into q_1 and q_2 in λ_1 and λ_2 but is not considered for now. To include simple demographic patterns a parameter μ representing a simple birth and death process is added here. For the sake of simplicity we assume that the overall population stays constant i.e. #births = #deaths. The overall population size is $N = 1 = S + I_1 + I_2 + I_{12}$. To make this model comparable to the ones analysed before one can study the behaviour of the model for $\mu \to 0$. This model's ecological dynamics for indistinguishable strains are:

$$\begin{aligned} N_0' &= -\beta(N_1 + 2qN_2)N_0 + wN_1 + \mu - \mu N_0 \\ N_1' &= +\beta(N_1 + 2qN_2)N_0 - wN_1 - c\beta(2I_1 I_2 + qN_2 N_1) + 2wN_2 - \mu N_1 \\ N_2' &= -2wN_2 + c\beta(2I_1 I_2 + qN_2 N_1) - \mu N_2, \end{aligned}$$

where $N_0 = S, N_1 = I_1 + I_2, N_2 = I_{12}$ Since there is no way to express these using ecological variables only, the model is not neutral. Anyway, now the stable equilibria will be assessed, in particular these with coexistence of two types. The disease free one again is obvious:

$$B_1 = (1, 0, 0, 0), \tag{4.3.11}$$

and the extinction of either one of the strains (which includes $I_{12}^s = 0$) leads to

steady states similar to those in the previously investigated models:

$$B_{2,3} = \{(\frac{1}{R_{01}}, 1 - \frac{1}{R_{01}}, 0, 0), (\frac{1}{R_{02}}, 0, 1 - \frac{1}{R_{02}}, 0)\} \tag{4.3.12}$$

with the strain specific reproduction number being $R_{0i} = \frac{\beta_i}{\mu + w_i}, i \in \{1, 2\}$. The steady state with coexistence of both types could be determined by successive elimination and substitution as done in the previous two paragraphs. This could be done by the introduction of auxiliary variables which represent all carriers of either serotype, as done by Lipsitch [37] and reviewed in detail by Weir [44]. Here, for convenience, the existence of this steady state will be inferred from the stability analysis of the disease free steady state and the steady states where only one type is present.

Stability analysis

The Jacobian matrix of the diamond-type model 4.3.10 is:

$$\begin{pmatrix} -(\lambda_1 + \lambda_2) - \mu & -\beta_1 S + w_1 & -\beta_2 S + w_2 & -qS(\beta_1 + \beta_2) \\ \lambda_1 & \beta_1 S - c_2\lambda_2 - w_1 - \mu & -c_2\beta_2 I_1 & w_2 + q\beta_1 S - c_2 q\beta_2 I_1 \\ \lambda_2 & -c_1\beta_1 I_2 & \beta_2 S - w_2 - \mu - c_1\lambda_1 & w_1 + q\beta_2 S - c_1 q\beta_1 I_2 \\ 0 & c_2\lambda_2 + c_1\beta_1 I_2 & c_1\lambda_1 + c_2\beta_2 I_1 & -\mu - (w_1 + w_2) + c_2\beta_2 q I_1 + c_1\beta_1 q I_2 \end{pmatrix} \tag{4.3.13}$$

Disease free equilibrium:

The Jacobian matrix in B_1 is:

$$J_1 = \begin{pmatrix} -\mu & -\beta_1 + w_1 & -\beta_2 + w_2 & q(\beta_1 + \beta_2) \\ 0 & \beta_1 - w_1 - \mu & 0 & w_2 + q\beta_1 \\ 0 & 0 & \beta_2 - w_2 - \mu & w_1 + q\beta_2 \\ 0 & 0 & 0 & -\mu - (w_1 + w_2) \end{pmatrix}. \tag{4.3.14}$$

Since the matrix 4.3.14 is an upper triangular matrix one can read off the Eigenvalues from the diagonal:

$$s_1 = -\mu, \quad s_2 = \beta_1 - w_1 - \mu, \quad s_3 = \beta_2 - w_2 - \mu, \quad s_4 = -\mu - (w_1 + w_2).$$

The Eigenvalues s_1 and s_4 are negative for the considered parameter space (i.e.

$\mu, w_{1,2} > 0$) and to ensure that $s_{2,3} \leq 0$

$$\frac{\beta_1}{\mu + w_1} = R_{01} \leq 1 \text{ and } \frac{\beta_2}{\mu + w_2} = R_{02} \leq 1$$

need to hold. So in case the reproduction numbers of both types are less then one or equal to one the disease free steady state is stable. Note that for asymptotic stability and therefore the convergence to B_1 the constraint $R_{0i} < 1,\ i = 1, 2$ is required.

Dominance of strain 1:

To derive the stability for B_2 an approach will be followed originally taken by Yidi Zhang [42] who splits the analysis into the equations for strain 1 and 2 and then makes use of the Perron-Frobenius theorem to derive the constrains on R_{01} and R_{02}. However, the model here is slightly different and the results cannot be inferred directly from Zhang's work (not considering strain specific loss of infectiousness rates in the SIS formulation of the model). The Jacobian matrix in B_2 is: $J_2 =$

$$\begin{pmatrix} -\left(1-\frac{1}{R_{01}}\right)\beta_1 - \mu & -\beta_1\frac{1}{R_{01}} + w_1 & -\beta_2\frac{1}{R_{01}} + w_2 & -q\frac{1}{R_{01}}(\beta_1+\beta_2) \\ \beta_1\left(1-\frac{1}{R_{01}}\right) & \beta_1\frac{1}{R_{01}} - w_1 - \mu & -c_2\beta_2\left(1-\frac{1}{R_{01}}\right) & w_2 + q\beta_1\frac{1}{R_{01}} - c_2q\beta_2\left(1-\frac{1}{R_{01}}\right) \\ 0 & 0 & \beta_2\frac{1}{R_{01}} - w_2 - \mu - c_1\beta_1\left(1-\frac{1}{R_{01}}\right) & w_1 + q\beta_2\frac{1}{R_{01}} \\ 0 & 0 & c_1\beta_1\left(1-\frac{1}{R_{01}}\right) + c_2\beta_2\left(1-\frac{1}{R_{01}}\right) & -\mu - (w_1+w_2) + c_2\beta_2 q\left(1-\frac{1}{R_{01}}\right) \end{pmatrix} \tag{4.3.15}$$

Since the matrix 4.3.15 is an upper triangular block matrix $\text{Det}(J_2 - sE_4) = \text{Det}(J_2^1 - sE_2) * \text{Det}(J_2^1 - sE_2)$ with E_n being the n-dimensional identity matrix,

$$J_2^1 = \begin{pmatrix} -\left(1-\frac{1}{R_{01}}\right)\beta_1 - \mu & -\mu \\ \beta_1\left(1-\frac{1}{R_{01}}\right) & 0 \end{pmatrix}$$

and

$$J_2^2 = \begin{pmatrix} \beta_2\frac{1}{R_{01}} - w_2 - \mu - c_1\beta_1\left(1-\frac{1}{R_{01}}\right) & w_1 + q\beta_2\frac{1}{R_{01}} \\ (c_1\beta_1 + c_2\beta_2)\left(1-\frac{1}{R_{01}}\right) & -\mu - (w_1+w_2) + c_2\beta_2 q\left(1-\frac{1}{R_{01}}\right) \end{pmatrix}$$

Note that J_2^1 has been reduced by making use of $\beta_1\frac{1}{R_{01}} = \mu + w_1$. The eigenvalues of J_2 are the same as these from J_2^1 and J_2^2 combined and the analysis can be split in two parts. J_2^1 here corresponds to the part of J_2 where the transmission dynamics of strain 1 are described. Since J_2^1 is two dimensional both eigenvalues are non-positive if the trace of J_2^1 is non-positive and the determinant is non-

negative.

$$\mathrm{Det}(J_2^1) = \mu\beta_1 \left(1 - \frac{1}{R_{01}}\right) \geq 0 \tag{4.3.16}$$

$$\mathrm{Trace}(J_2^1) = -\left(1 - \frac{1}{R_{01}}\right)\beta_1 - \mu \leq 0 \tag{4.3.17}$$

From equation 4.3.16 follows the need for

$$R_{01} \geq 1 \tag{4.3.18}$$

which is then sufficient to fulfil equation 4.3.17 in the considered parameter space as well. For the second part of the stability analysis for B_2, J_2^2 can be split into

$$J_2^{2,1} = \begin{pmatrix} -w_2 - \mu - c_1\beta_1\left(1 - \frac{1}{R_{01}}\right) & w_1 \\ c_1\beta_1\left(1 - \frac{1}{R_{01}}\right) & -\mu - (w_1 + w_2) \end{pmatrix}$$

and

$$J_2^{2,2} = \frac{1}{R_{01}} \begin{pmatrix} 1 & q \\ c_2(R_{01} - 1) & c_2 q(R_{01} - 1) \end{pmatrix}$$

with $J_2^2 = J_2^{2,1} + \beta_2 J_2^{2,2}$. Since, given $R_{01} \geq 1$, $J_2^{2,1}$ is a quasi-positive matrix and $J_2^{2,2}$ a positive matrix the Perron-Frobenius theorem [45] is applicable. Hence, the spectral bound $s(J_2^2)$ (the largest real part of all eigenvalues of J_2^2) is itself an eigenvalue of J_2^2. Using the determinant and the trace as before one can easily show, that $J_2^{2,1}$ ($J_2^2(\beta_2 = 0)$) has two negative eigenvalues (i.e. $s(J_2^2(\beta_2 = 0)) < 0$). Since the spectral bound is monotonously increasing for increasing β_2, $\beta_2 \leq \beta_2^*$ with $\mathrm{Det}(J_2^2(\beta_2 = \beta_2^*)) = 0$ ensures negative eigenvalues for J_2^2:

$$\begin{aligned} \beta_2^* \Bigg[\frac{1}{R_{01}}(-\mu - w_1 - w_2) + \left(-\mu - w_2 - c_1\beta_1\left(1 - \frac{1}{R_{01}}\right)\right) c_2 q \left(1 - \frac{1}{R_{01}}\right) \\ - \left(c_1\beta_1 q \frac{1}{R_{01}} + c_2 w_2\right)\left(1 - \frac{1}{R_{01}}\right) \Bigg] \\ = \left(w_2 + \mu + c_1(w_1 + \mu)(R_{01} - 1)\right)(-\mu - w_2) + w_1(\mu + w_2) \end{aligned}$$

therefore the constraint on R_{02} is:

$$R_{02} \leq \frac{\beta_2^*}{w_2+\mu} \tag{4.3.19}$$

$$= R_{01}\frac{1+\frac{w_1}{w_2+\mu}+(R_{01}-1)\frac{w_1+\mu}{w_2+\mu}c_1}{1+\frac{w_1}{w_2+\mu}+(R_{01}-1)\left(\frac{w_1+\mu}{w_2+\mu}c_1q(c_2(R_{01}-1)+1)+c_2(q+\frac{w_1}{w_2+\mu})\right)}. \tag{4.3.20}$$

Note that for $R_{01} \geq 1$ the right hand side is positive and therefore allows a meaningful constraint.

Dominance of strain 2:

Since the Model is symmetric the stability for B_3 can be derived exactly as for B_2. The constraints corresponding to 4.3.18 and 4.3.20 are:

$$R_{02} \geq 1 \tag{4.3.21}$$

and

$$R_{01} \leq R_{02}\frac{1+\frac{w_2}{w_1+\mu}+(R_{02}-1)\frac{w_2+\mu}{w_1+\mu}c_2}{1+\frac{w_2}{w_1+\mu}+(R_{02}-1)\left(\frac{w_2+\mu}{w_1+\mu}c_2q(c_1(R_{02}-1)+1)+c_1(q+\frac{w_2}{w_1+\mu})\right)}. \tag{4.3.22}$$

Inferring the domain for coexistence of both strains

The previous analysis has determined the domain for R_{01} and R_{02} where the steady states B_1, B_2 and B_3 are insensitive to small perturbations. In particular the results show that a steady state where one strain becomes extinct is only stable if this strain has a reproduction number bigger or equal to one and the second strain's reproduction number is less and sufficiently small so as not to invade the population. This means that in case the latter does not hold the second strain does (at least temporarily) increase in prevalence. Therefore it is likely (see sections 4.3.2 and 4.3.3), that the diamond-type model has stable coexistence of both strains in the domain of $R_{01,02}$ where $B_{1,2,3}$ are not stable. This domain is $C = C_1 \cap C_2 \cap C_3$ with

$$C_1 = \left\{ (R_{01},R_{02}) \middle| R_{01} \geq \frac{R_{02}\left(1+\frac{w_2}{w_1+\mu}+(R_{02}-1)\frac{w_2+\mu}{w_1+\mu}c_2\right)}{1+\frac{w_2}{w_1+\mu}+(R_{02}-1)\left(\frac{w_2+\mu}{w_1+\mu}c_2q(c_1(R_{02}-1)+1)+c_1(q+\frac{w_2}{w_2+\mu})\right)} \right\} \tag{4.3.23}$$

$$C_2 = \left\{ (R_{01}, R_{02}) \middle| R_{02} \geq \frac{R_{01}\left(1 + \frac{w_1}{w_2+\mu} + (R_{01}-1)\frac{w_1+\mu}{w_2+\mu}c_1\right)}{1 + \frac{w_1}{w_2+\mu} + (R_{01}-1)\left(\frac{w_1+\mu}{w_2+\mu}c_1 q(c_2(R_{01}-1)+1) + c_2(q + \frac{w_1}{w_1+\mu})\right)} \right\} \tag{4.3.24}$$

and

$$C_3 = \{(R_{01}, R_{02}) | R_{01}, R_{02} \geq 1\} \tag{4.3.25}$$

To test the existence of steady states of coexistence in C numerically the diamond-type model was implemented and numerically solved using the explicit Euler method with a time-step of 1. For a fixed set of parameters $w_{1,2}, \mu, c_{1,2}$ and q and initial values $S, I_1, I_2 > 0$ and $0 \leq I_{12} = 1 - S - I_1 - I_2$ 1000 model simulation were run. Each simulation sampled R_1 and R_2 from a uniform distribution between 0 and 5 and was run until the sum of the Euclidean differences of S, I_1, I_2 and I_{12} at times t and $t-1$ fell under an error threshold of $1 * 10^{-7}$ (the model was then considered to be in a steady state). A strain was considered extinct if at steady state less then 5‰(other thresholds yielded similar results) of the population were infected. The simulations confirmed the hypothesis of C being the domain which grants a stable steady state of coexisting strains (see Figure 4.2). Furthermore, by its design the implemented model suggests global stability of the diamond type model on C as well.

While the size of C is relatively robust against different assumptions of the loss of infectiousness (w_i) and mortality (μ) it is mainly dependent on the competition parameters c_1, c_2 as well as q (See Figure 4.3). One also observes that the higher the competition between the strains (recall that high competition in this model is reflected by low values of $c_{1,2}$) the smaller the parameter space which allows for coexistence (in particular in case $c_1 = c_2 = 0$ coexistence is restricted to $R_{01} = R_{02}$ - see equations 4.3.23 and 4.3.24 as well).

4.3.5 The issue of coexistence

While the diamond-type model provides a reasonable framework to describe co-existence of two strains (for example vaccine-types and non vaccine-types) and ensures stable coexistence for a broad range of parameter values it lacks neutrality [41]. This makes it unsuitable for exploring the underlying mechanism of coexistence since it promotes them artificially. In particular the explored stability of the diamond-type model is in disagreement with the necessity of equally fit

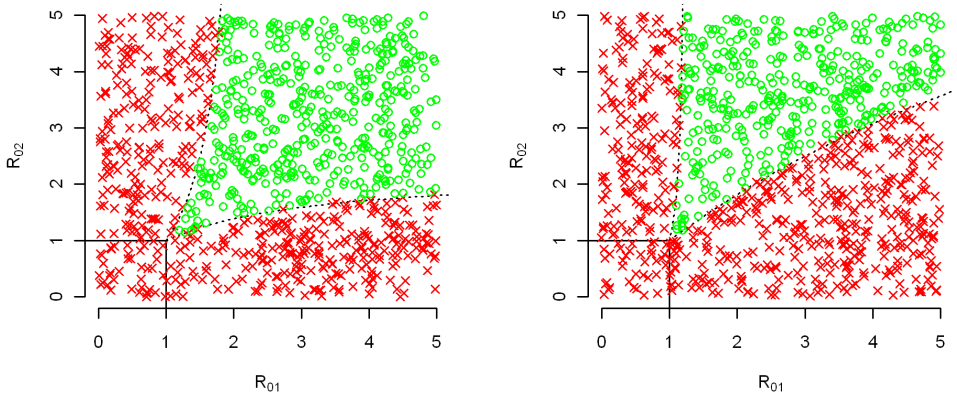


Figure 4.2: Numerical model results for the test of stable coexistence. The dashed line marks the domain given by the intersection of 4.3.23, 4.3.24 and 4.3.25. Model simulations which resulted in coexistence of both strains are represented by green circles and the extinction of at least one strain by red crosses. The figure consists of 1000 simulations with uniformly distributed R_{01} and R_{02}, the other parameter values are: $w_1 = 1/60, w_2 = 1/60, \mu = 1/1000, c_1 = 0.9, c_2 = 0.9, q = 1/2$ (left), $w_1 = 1/6, w_2 = 1/6, \mu = 1/10000, c_1 = 0.8, c_2 = 0.1, q = 1$ (right) and the initial values: $(S, I_1, I_2, I_{12}) = (0.3, 0.3, 0.3, 0.1)$.

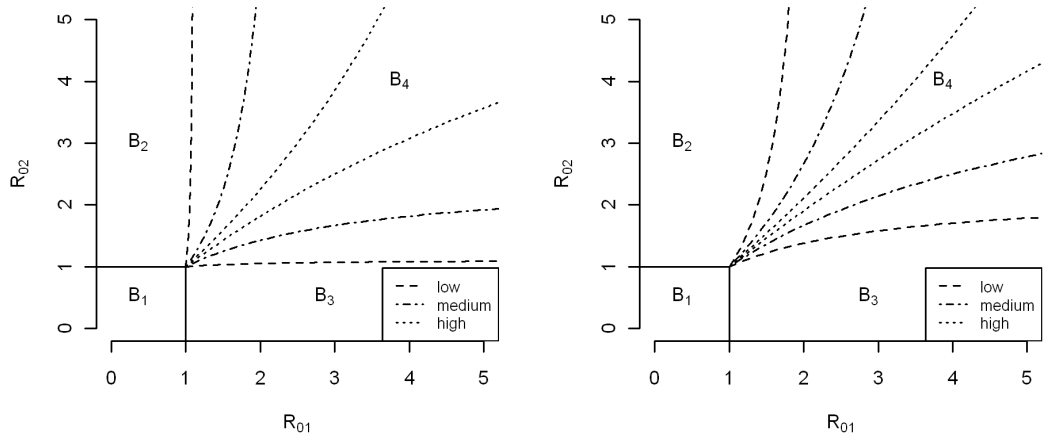
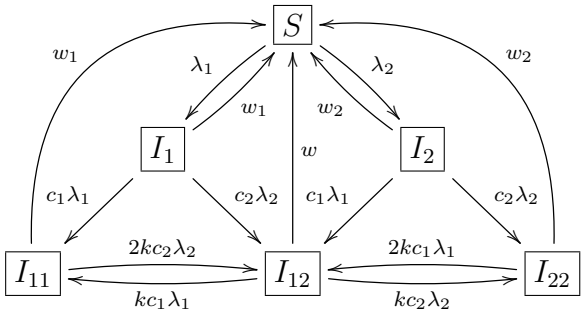


Figure 4.3: Domains for R_{01} and R_{02} which corresponds to the domain of stability of the steady states $B_1 \dots B_4$. Parameter values are: $w_1 = w_2 = 1/60$, $\mu = 1/1000$, $c_1 = c_2$, $q = 1$(left), $q = 0.5$(right) for three scenarios: low ($c_1 = 0.9$), medium ($c_1 = 0.4$) and high ($c_1 = 0.1$) competition between the strains.

strains being able to coexist at different prevalence levels. Hence forecasts relying on model predictions of a steady state should be treated with great care, however, in the absence of feasible alternatives they still can provide a helpful best guess.

Assessing the mechanisms which promote coexistence in neutral models Colijn et al. [46] were able to identify a few essential factors. In essence they explored variants of the diamond-type model including extra compartments for individuals infected twice with the same strain. The basic model is described as



where k represents the probability for "knocking off" one strain when exposed to another. When intra-specific competition exhibits inter-specific competition in a classic Lotka-Voltera model coexistence is promoted. The authors transferred this to their models and found the same mechanism to strongly support coexistence. In particular this means, that if a person infected with strain 1 has a reduced chance to get infected with strain 1 opposed to strain 2 and vice versa, coexistence is promoted. Serotype specific immunity does present a mechanism which has been observed in longitudinal carriage studies (see section 2.4.3) and could be a biological ground for intra-specific competition. This will become important in chapter 6 of this thesis.

Summary

The transmission dynamics of the pneumococcus were discussed in this chapter with a focus on the role of competition for the spread, and especially, for coexistence of the different serotypes. The means of acquisition, clearance, replacement and competition were discussed and the concept of a neutral model was discussed. Three groups of deterministic models including different approaches for incorporating coexistence were analysed for their steady state behaviour in the light of model neutrality. The two neutral model formulations did allow co-

existence of two types only for a limited parameter space (e.g. if the types are similarly fit: $R_{01} \approx R_{02}$) which is likely to prove too restrictive to use it for model fitting in the case of *S. pneumoniae*. Further neither models incorporated a compartment which allows colonisation with both serotypes at once as was observed for the pneumococcus. The non neutral model has this extra compartment and further benefits from a large parameter space where coexistence is possible. However, through the model structure, this approach suffers from the caveat of artificially increasing coexistence in absence of any biologically plausible mechanism underpinning these dynamics. This issue could lead to biased estimates of pneumococcal carriage if used for extrapolation and needs to be kept in mind when employing such model structures to gain insights as to the likely impact of vaccination.

The subsequent chapter will review a modelling approach for predicting the long term impact of introducing PCV7 into the childhood immunisation scheme of England and Wales and its evaluation when applied to the observed changes in disease. The feasibility of this model to mirror the epidemiology observed in Scotland will be studied in the perspective of extending the model to predict the impact of higher valency vaccine formulations. The deterministic model will be converted into a corresponding individual based model in order to cope with the complex vaccination uptake following the introduction of PCV13 in England and Wales. The caveats for this transfer will be discussed and insight gained will be used when evaluating the possible role of immunity in the transmission dynamics for inducing competition and coexistence.

Bibliography

[1] Melegaro A, Gay NJ, Medley GF (2004) Estimating the transmission parameters of pneumococcal carriage in households. Epidemiol Infect 132: 433–441.

[2] Hoegberg L, Geli P, Ringberg H, Melander E, Lipsitch M, et al. (2007) Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. Journal of Clinical Microbiology 45: 948–952.

[3] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, et al. (2010) Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC infectious diseases 10: 90.

[4] Melegaro A, Choi Y, Pebody R, Gay N (2007) Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. American journal of epidemiology 166: 228–35.

[5] Weinberger DM, Trzccinski K, Lu YJ, Bogaert D, Brandes A, et al. (2009) Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathogens 5: e1000476.

[6] Sleeman KL, Griffiths D, Shackley F, Diggle L, Gupta S, et al. (2006) Capsular serotype-specific attack rates and duration of carriage of *Streptococcus pneumoniae* in a population of children. Journal of Infectious Diseases 194: 682–688.

[7] Dagan R, Givon-Lavi N, Zamir O, Sikuler-Cohen M, Guy L, et al. (2002) Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after ad-

ministration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. Journal of Infectious Diseases 185: 927–936.

[8] O'Brien KL, Millar EV, Zell ER, Bronsdon M, Weatherholtz R, et al. (2007) Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. The Journal of infectious diseases 196: 1211–20.

[9] Gray BM, Converse GM, Dillon HC (1980) Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. Journal of Infectious Diseases 142: 923–33.

[10] Kuster SP, Tuite AR, Kwong JC, McGeer A, Fisman DN (2011) Evaluation of Coseasonality of Influenza and Invasive Pneumococcal Disease: Results from Prospective Surveillance. PLoS medicine 8: e1001042.

[11] Gray BM, Turner ME, Dillon HC (1982) Epidemiologic studies of *Streptococcus pneumoniae* in infants. The effects of season and age on pneumococcal acquisition and carriage in the first 24 months of life. In: American Journal of Epidemiology. volume 116, pp. 692–703.

[12] Syrjänen RK, Kilpi TM, Kaijalainen TH, Herva EE, Takala AK (2001) Nasopharyngeal carriage of *Streptococcus pneumoniae* in Finnish children younger than 2 years old. The Journal of infectious diseases 184: 451–9.

[13] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[14] Lakshman R, Murdoch C, Race G, Burkinshaw R, Shaw L, et al. (2003) Pneumococcal nasopharyngeal carriage in children following heptavalent pneumococcal conjugate vaccination in infancy. Archives of Disease in Childhood 88: 211–214.

[15] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[16] Huang SS, Hinrichsen VL, Stevenson AE, Rifas-Shiman SL, Kleinman K, et al. (2009) Continued impact of pneumococcal conjugate vaccine on car-

riage in young children. Pediatrics 124: e1–11.

[17] Vestrheim DF, Hø iby EA, Aaberge IS, Caugant DA (2010) Impact of a pneumococcal conjugate vaccination program on carriage among children in Norway. Clinical and Vaccine Immunology 17: 325–334.

[18] Spijkerman J, van Gils EJM, Veenhoven RH, Hak E, Yzerman EPF, et al. (2011) Carriage of *Streptococcus pneumoniae* 3 Years after Start of Vaccination Program, the Netherlands. Emerging infectious diseases 17: 584–591.

[19] Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease after pneumococcal vaccination. Lancet 6736: 1–12.

[20] Auranen K, Mehtala J, Tanskanen A, Kaltoft MS (2010) Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. American journal of epidemiology 171: 169–176.

[21] Weinberger DM, Harboe ZB, Flasche S, Scott JA, Lipsitch M (2011) Prediction of Serotypes Causing Invasive Pneumococcal Disease in Unvaccinated and Vaccinated Populations. Epidemiology 22: 199–207.

[22] Gupta S, Maiden MC, Feavers IM, Nee S, May RM, et al. (1996) The maintenance of strain structure in populations of recombining infectious agents. Nature medicine 2: 437–42.

[23] Gupta S, Swinton J, Anderson RM (1994) Theoretical studies of the effects of heterogeneity in the parasite population on the transmission dynamics of malaria. Proceedings Biological sciences / The Royal Society 256: 231–8.

[24] White LJ, Cox MJ, Medley GF (1998) Cross immunity and vaccination against multiple microparasite strains. IMA journal of mathematics applied in medicine and biology 15: 211–33.

[25] Rozenbaum MH, Sanders EAM, van Hoek AJ, Jansen AGSC, van der Ende A, et al. (2010) Cost effectiveness of pneumococcal vaccination among Dutch infants: economic analysis of the seven valent pneumococcal conjugated vaccine and forecast for the 10 valent and 13 valent vaccines. BMJ 340: c2509.

[26] Smith KJ, Wateska AR, Nowalk MP, Raymund M, Nuorti JP, et al. (2012) Cost-effectiveness of adult vaccination strategies using pneumococcal conju-

gate vaccine compared with pneumococcal polysaccharide vaccine. JAMA : the journal of the American Medical Association 307: 804–12.

[27] Andersson M, Ekdahl K, Mölstad S, Persson K, Hansson HB, et al. (2005) Modelling the spread of penicillin-resistant *Streptococcus pneumoniae* in day-care and evaluation of intervention. Statistics in medicine 24: 3593–607.

[28] Sutton KL, Banks HT, Castillo-Chavez C (2008) Estimation of invasive pneumococcal disease dynamics parameters and the impact of conjugate vaccination in Australia. Mathematical biosciences and engineering : MBE 5: 175–204.

[29] Lipsitch M (2001) Measuring and interpreting associations between antibiotic use and penicillin resistance in *Streptococcus pneumoniae*. Clinical infectious diseases 32: 1044.

[30] McCormick AW, Whitney CG, Farley MM, Lynfield R, Harrison LH, et al. (2003) Geographic diversity and temporal trends of antimicrobial resistance in *Streptococcus pneumoniae* in the United States. Nature medicine 9: 424–30.

[31] Temime L, Boëlle PY, Courvalin P, Guillemot D (2003) Bacterial resistance to penicillin G by decreased affinity of penicillin-binding proteins: a mathematical model. Emerging infectious diseases 9: 411–7.

[32] Temime L, Guillemot D, Boëlle PY (2004) Short- and long-term effects of pneumococcal conjugate vaccination of children on penicillin resistance. Antimicrobial agents and chemotherapy 48: 2206–13.

[33] Temime L, Boëlle PY, Valleron aJ, Guillemot D (2005) Penicillin-resistant pneumococcal meningitis: high antibiotic exposure impedes new vaccine protection. Epidemiology and infection 133: 493–501.

[34] Wang YC, Lipsitch M (2006) Upgrading antibiotic use within a class: tradeoff between resistance and treatment success. Proceedings of the National Academy of Sciences of the United States of America 103: 9655–60.

[35] Opatowski L, Temime L, Varon E, Leclercq R, Leclerc R, et al. (2008) Antibiotic innovation may contribute to slowing the dissemination of multiresistant *Streptococcus pneumoniae*: the example of ketolides. PloS one 3: e2089.

[36] Erästö P, Hoti F, Granat SM, Mia Z, Mäkelä PH, et al. (2010) Modelling multi-type transmission of pneumococcal carriage in Bangladeshi families. Epidemiology and infection 138: 861–72.

[37] Lipsitch M (1997) Vaccination against colonizing bacteria with multiple serotypes. Proceedings of the National Academy of Sciences 94: 6571–6576.

[38] Temime L, Boelle PY, Opatowski L, Guillemot D (2008) Impact of capsular switch on invasive pneumococcal disease incidence in a vaccinated population. PloS one 3: e3244.

[39] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[40] Choi YH, Jit M, Flasche S, Gay N, Miller E (2012) Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. PLoS ONE 7: e39927.

[41] Lipsitch M, Colijn C, Cohen T, Hanage W, Fraser C (2009) No coexistence for free: neutral null models for multistrain pathogens. Epidemics 1: 2.

[42] Zhang Y (2006) Mathematische Modelle zur analyse der Koexistenz zweier konkurierender Serotypen. Ph.D. thesis, Unviversity Thuebingen.

[43] Zhang Y, Auranen K, Eichner M (2004) The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. Epidemiology and infection 132: 1073–81.

[44] Weir A (2009) Modelling the impact of vaccination and competition on pneumococcal carriage and disease in Scotland. Ph.D. thesis, University of Strathclyde.

[45] Thieme H (2006) Perron-Frobenius theory of positive matrices. In: Mathematics in Population Biology, Princton: Princeton University Press. p. 447.

[46] Colijn C, Cohen T, Fraser C, Hanage W, Goldstein E, et al. (2010) What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae*? Journal of the Royal Society, Interface / the Royal Society 7: 905–19.

5 A model to predict the impact of vaccination

Outline

The complex epidemiology of *S. pneumoniae*, both before and in response to the introduction of the 7-valent vaccine in the UK and elsewhere, has been described in chapter 3. The substantial decrease in vaccine type IPD in both unvaccinated and vaccinated individuals, through a combination of a protective effect against pneumococcal disease and direct and indirect prevention of pneumococcal colonisation by the vaccine, was subsequently followed by a marked increase in NVT IPD caused by serotype replacement. In chapter 4 possible model structures to describe the basic transmission dynamics of the pneumococcus were studied. The "diamond type model" which provides the most promising approach to mirror coexistence of multiple types has the caveat of providing coexistence through the model structure itself rather than a biologically plausible mechanism.

Following the success of PCV7 in the United States dynamic modelling was used to infer possible changes in IPD in the England and Wales [1, 2] and to assist the decision on implementing a similar program. This approach employed a diamond type model similar to the one discussed in section 4.3.4. In this chapter these manuscripts will be reviewed and the techniques used will be applied to data on vaccine uptake and invasive pneumococcal disease in Scotland. PCV13 has already been introduced in Scotland; therefore studying the possible impact of PCV7 might seem uncalled for. However, this could provide an estimate of the long term impact of PCV7 as introduction of PCV13 made such an assessment by other means impossible. Estimation of the long term impact of PCV7 is necessary if one ought to appropriately study the incremental benefit of the introduction of

PCV13. Furthermore knowledge of the impact of PCV7 offers a basis to study the possible long term impact of PCV13 in Scotland. Finally, the model is transferred into a corresponding individual based framework which is more flexible and in particular allows inclusion of a more complex structure of protection following vaccination after the switch in policy from PCV7 to PCV13.

5.1 Predicting the impact of PCV7 in E&W

A model structurally similar to the one analysed in chapter 4.3.4 was set up by Melegaro and colleagues [1] to predict the possible impact of the introduction of PCV7 into the UK childhood immunisation scheme following the huge impact reported from the US [3]. In this model serotypes included in the vaccine formulation were treated as one type which was competing with another type which represents all non-vaccine serotypes. To allow for imperfect vaccination protection the authors employed two parallel diamond type models where individuals can switch between those according to vaccination rates and waning vaccine protection. The model was run until it reached a steady state and age-stratified values for the FOI and the invasiveness of VTs and NVTs were estimated from pre-vaccination carriage and disease data. The key model parameters were estimated from the dynamics of IPD in the US in the post-vaccination era by minimising the Poisson deviance. Finally the model was re-parametrized with the UK population estimates and pre-vaccination IPD data to forecast the effect in the UK.

Vaccine-type IPD in England and Wales was estimated to decline quickly and to vanish about 10 years after the introduction of PCV7 due to its direct and indirect (herd) protection. The protection against non-vaccine serotype acquisition through vaccine type colonisation was estimated to be only about 15%. This determines the amount of competition between VT and NVT and is therefore crucial for the prediction of the magnitude of NVT replacement of which there was not much evidence in the US at that time (which was partly down to an increasing number of outpatients included in the US IPD data, see section 3.2 and [4]). In total the model predicted a 63% reduction of IPD cases in children less then 5 years of age and a 35% reduction in the rest of the population.

Shortly after introduction of PCV in the UK in September 2006 it became increasingly clear that, even though the decline of VT IPD in the UK somewhat

mirrored the observations from the US and indeed the model predictions, the increase in NVT disease was much more pronounced than expected (compare chapter 3.2). Only some of the discrepancies disappeared when accounting for pre-vaccination trends in the data from E&W and at the same time reducing the bias from changing case ascertainment in the post-vaccination era in the US by including only hospitalised cases.

This showed the need for re-parameterisation of the model to UK post-vaccination data so that the long term effects of PCV7 can be predicted more accurately. Choi et al. adapted the model to incorporate the observed vaccine uptake in E&W in which another copy of the diamond model was introduced into the framework to incorporate the state of partial protection of individuals who received only one dose of the vaccine until completion of their first year of life [2]. A model diagram is presented in figure 5.1. Also the implementation of the mixing patterns was changed to correspond to those observed in a contact survey conducted in the UK and other European countries in 2006 [5]. As previously the model was fitted to pre-vaccination carriage prevalence obtained from a longitudinal study in Hertfordshire (see [6, 7] and 3.1.3) and the invasiveness of VTs and NVTs was determined from observed pre-vaccination IPD levels.

Assuming that the invasiveness of the grouped serotypes in VT and NVT did not change the model was then fitted to post-vaccination IPD of E&W until the epidemiological year 2008/09. The IPD data was adjusted for increased ascertainment (see section 3.1.1) and excluded serotype 1 because of its volatility. The assumption of constant invasiveness could be flawed if within the group of NVT strains only a few serotypes with outstanding invasiveness were responsible for most of the replacement. However, if the rank order of carriage prevalence is determined by pathogen specific characteristics (for which there is some evidence [8, 9]) the assumption will hold true in the long term, although it might be prone to some stochastic variation.

The best fitting model estimated that individuals colonised with VT are 96% [82%-100%] less likely to acquire a NVT compared to uncolonised individuals. This finding is similar to the one estimated directly from a longitudinal dataset for pneumococcal carriage in three Danish day-care centres [10]. Auranen et al. found that the rate of acquisition of infection by a pneumococcal serotype in uninfected children (pooled over all serotypes) was reduced by 91%[85%-95%] when they were infected. Through this high amount of competition in the model

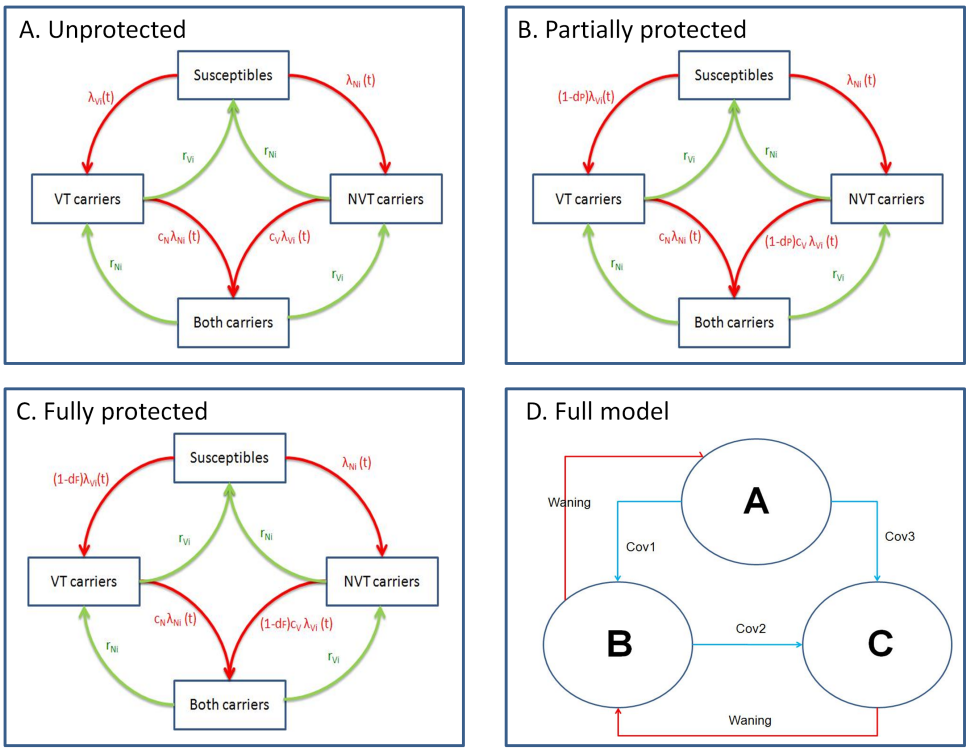


Figure 5.1: Flow diagram of the transmission dynamics employed in the model. Sus: susceptibles, VT: PCV7 type carriers, NVT: non-PCV7 type carriers, Both: both carrying VT and NVT. A) Pre-PCV7 or Unprotected group, B) Partially Protected, C) Fully Protected, and D) Movements between three groups (A, B and C) through vaccination and waning. The parameters $r_i, i \in \{Ni, Vi\}$ are age-dependent recovery rates for NVT and VT, $\lambda_i, i \in \{Ni, Vi\}$ are forces of infection, and c_{Ni} and c_{Vi} are competition parameters. The vaccine protection amongst the population is defined by the vaccine coverage rates $cov_i, i \in \{1, 2, 3\}$. The vaccine efficacy against carriage for partially and fully protected groups are denoted as d_P and d_F respectively [2].

the long term prediction showed almost complete replacement of VT carriage with NVTs and so the impact of PCV7 on overall IPD was lower; elimination of VT was predicted to occur within 10 years after implementation of the program and NVT IPD to increase by 90%[63%-111%] which in the long term results in an overall reduction of IPD cases of only 9%[2%-22%] and therefore to be considerably less beneficial than estimated from US data.

5.2 Predicting the impact of PCV7 in Scotland

Following the assessment of the impact of PCV7 in Scotland (see section 3.1.2) this section aims to evaluate the feasibility of the model from Choi et al. [2] for describing the Scottish IPD data and predicting the likely long term population effect of PCV7. This would give an opportunity to unravel the effects of PCV7 and PCV13 after the introduction of PCV13 in 2010. Also this offers the perspective of further adapting the model to estimate the potential impact of higher valency vaccine formulations.

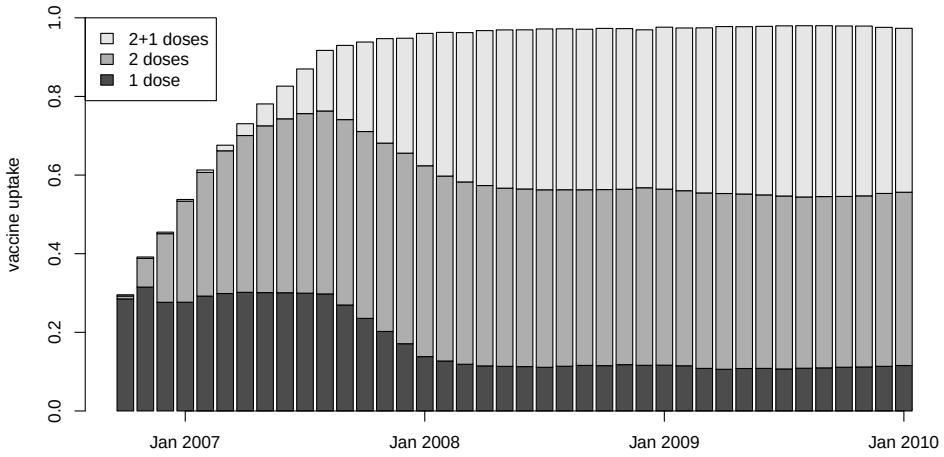


Figure 5.2: Vaccine (PCV7) uptake in 2-24 month old Scottish children in the post PCV7 introduction era.

5.2.1 Methods

Data & parameterisation Data on invasive pneumococcal disease obtained from blood or cerebrospinal fluid from the Scottish national surveillance from 2000/01 to 2009/10 were extracted from the SPIDER database (see chapter 3.1.2). The range of the age groups were chosen achieve equally distributed numbers of cases amongst the age group providing about 100 cases per year in each group in order to minimise differences in stochastic effects. The age groups were: 0-4y, 5-39y, 40-49y, 50-59y, 60-69y, 70-79y, 80+y. For estimates on pneumococcal colonisation data from a pre-vaccination study in Hertfordshire, England, were employed [6], because no data in nasopharyngeal colonisation from Scotland is available for the pre vaccination era. Data on vaccine uptake in Scotland after introduction of PCV7 was obtained from the Scottish Immunization Record System (SIRS). Mixing patterns where implemented according to a survey by Mossong et al. interviewing people about their day to day two-way conversational contact behaviour which is assumed to reflect events mainly responsible for transmission of various pathogens including the pneumococcus [5]. Age-dependent duration of carriage was assumed to be similar to these employed in Melegaro et al. [1] and the remaining parameters including duration of vaccine protection and vaccine effectiveness against invasive disease given colonisation were obtained from Choi et al. [2].

Model The program code for the model was obtained from the authors of Choi et al. [2] which is an extension of the diamond-type model described in section 4.3.4 to include full and partial vaccine protection (see figure 5.1). A fully detailed description of the model can be found in Choi et al. [2]. In brief: in the model people change respective compartments according to their age dependent probability of being vaccinated. This is determined by vaccine uptake data. Individuals are considered fully protected from either two or more doses or by receiving one dose above the age of 1 (catch up campaign). Everyone else who received one dose is considered partially protected and is therefore assumed to be only half as protected against acquisition of infection as a fully protected person.

Data analysis As presented in section 3.1.2 the interpretation of post vaccination IPD in Scotland was hindered by increasing incidence prior to the introduction of PCV7. Hence, the analysis here is presented in three scenarios. For

scenarios 1 and 3 data on invasive disease was adjusted for what is assumed to be an effect of increasing ascertainment; the Poisson regression model from section 3.1.2 was employed to estimate the yearly increase in overall IPD in the absence of vaccination. The IPD data in the epidemiological year 2009/10 was hereby assumed to represent the year when ascertainment was complete and the number of cases (k) in the years $t = 1, \ldots, 8$ (corresponding to the epidemiological years 2000/01, ..., 2008/09) where inflated:

$$k_t^{inflated} = k_t \cdot s^{9-t}$$

Here s represents the pre-vaccination slope estimated by the regression model. Further, in scenario 1 serotype 1 disease was excluded for reasons discussed earlier (see chapter 2.6.2). Overall scenario 1 represents the analysis fitting to fully adjusted data, scenario 2 to unadjusted data (including serotype 1) and in scenario 3 the IPD data was adjusted for trends but included serotype 1 IPD.

In contrast to Choi et al. model fitting was done through Markov Chain Monte Carlo methods and the program code was adapted accordingly. A burn-in period of 10,000 simulations was allowed for and a total chain length of 100,000 was considered sufficient. No formal assessment of convergence was employed. The Poisson log-likelihood of the model estimates (λ^c for carriage and λ^I for IPD), given one year of pre vaccination carriage data (j) and 4 years of post vaccination IPD (k) data, is:

$$\begin{aligned} LL(\lambda^I, \lambda^c, k, j) = & \sum_{a=1}^{7} \sum_{y=2}^{5} \left[-\lambda_{a,y}^I + k_{a,y}\ log(\lambda_{a,y}^I) - log(k_{a,y}!) \right] + \\ & \sum_{b=1}^{10} \left[-\lambda_{b,1}^c + j_{b,1}\ log(\lambda_{b,1}^c) - log(j_{b,1}!) \right]. \end{aligned}$$

The model estimates are assumed to be the mean of a Poisson distributed variable. The variables a and b represent the 7 and 10 age groups employed for invasive disease and nasopharyngeal carriage respectively and $y \in \{1, \ldots, 5\}$ the pre vaccination ($y = 1$) and the 4 post vaccination years. The log-likelihood was maximised through a Metropolis Hastings algorithm. For computational efficiency $log(k!)$ was omitted since there was no interest in the likelihood itself but rather in the likelihood ratio to calculate the acceptance rate. The fitting parameters (competition ($c_N = c_V$), the vaccine induced degree of protection against carriage (ξ) and the age-dependent probability of colonisation given a potentially

infective contact stratified by VT and NVT and three age groups: 0-1y, 2-4y and 5+ ($q_1 \dots q_6$)) where estimated through a random walk with a normal proposal distribution. The standard deviation of the respective proposal distributions were adjusted to achieve a common acceptance rate of about 23.4% (between 15% and 50%) to achieve optimal mixing properties of the chain [11]. All parameters we bounded to be between 0 and 1 which was achieved by penalising the likelihood if any parameter exceeds these bounds. As Choi et al found, the model is insensitive to the competition parameter representing the acquisition of a VT given NVT carriage and therefore both competition parameters where assumed to be the same. Uninformative priors where employed. The model was programmed in Matlab and data analysis was performed R.

5.2.2 Results

After introduction in September 2006 PCV7 coverage in Scotland quickly reached levels exceeding 95% in the 2-24 month olds (see figure 5.2). Less than 10% of eligible individuals received less then 3 doses by the age of 18 months.

The Markov Chains for all three scenarios converged (figures 5.3, 5.7 and 5.11) to unimodal posterior distributions (see the diagonals of figures 5.4, 5.8 and 5.12). Both convergence and posterior distributions were found to be insensitive to the choice of starting points. There was some correlation between the parameters representing the age-dependent risk of an infectious contact as well as between vaccine effectiveness against carriage and the competition (see figures 5.4, 5.8 and 5.12). This was consistent amongst all scenarios and suggests that re-parameterisation of the model could lead to quicker convergence and more efficient mixing of the Markov chain. However the brute force approach taken here worked fine for this rather well behaved and low dimensional problem.

Accounting for increasing ascertainment in the pre-vaccination era and neglecting ST1 disease (scenario 1) the model did fit the pre vaccination carriage prevalence well (see figure 5.5). Parameter estimates are presented in table 5.1.Individuals colonised with a VT were estimated to be about half as likely to acquire a NVT than uncolonised individuals; the reduction in acquisition given colonisation (1-competition) was estimated to be 44%, 95% credible interval: [35, 55]. The risk of an infectious contact was estimated to be highest in the very young population and to be slightly bigger for infection with VT than for infection with NVT.

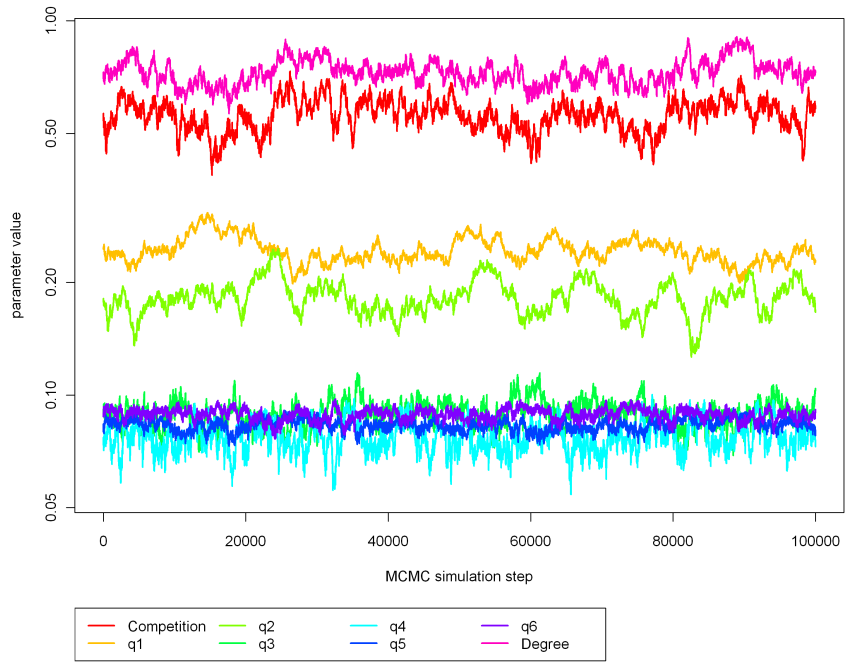


Figure 5.3: Scenario 1 (trend-adjusted; exclusion of ST1): Markov Chains for fitting the parameters competition, degree and the six q-values.

The risks for an infectious contact with VT in 0-1y, 2-4y and 5+ (q_1, q_2, q_3) were 0.243 [0.214, 0.275], 0.090 [0.076, 0.104] and 0.083 [0.077, 0.089]. The risks for an infectious contact with NVT (q_4, q_5, q_6) were 0.182 [0.147, 0.213], 0.076 [0.063, 0.092] and 0.089 [0.084, 0.095] respectively. As in Choi et al. vaccine effectiveness against disease was assumed to be 100%. Vaccine effectiveness against carriage was estimated to be 72.3% [62.6, 82.7]. Obtained case:carrier ratios were highest in children younger than 5 years of age and in the elderly (70+). While in children the invasiveness of VT was much higher than the invasiveness of NVT, adults were estimated to be slightly more susceptible to disease from NVT.

The model fits the data on overall IPD in Scotland well (figure 5.6), however stratified by age one finds a lack of agreement especially in the age groups 5-39y and 80+. The Scottish IPD data in the 5-39y old population suggests less VT replacement than estimated from the model while the data in people aged 80 years or older suggests much more serotype replacement then predicted by the model.

Including serotype 1 in the analysis and not adjusting for trends in the pre-vaccination era (scenario 2) yielded slightly different parameter estimates (see figure 5.8 and table 5.1). The carriage data is again fitted well (see figure 5.9).

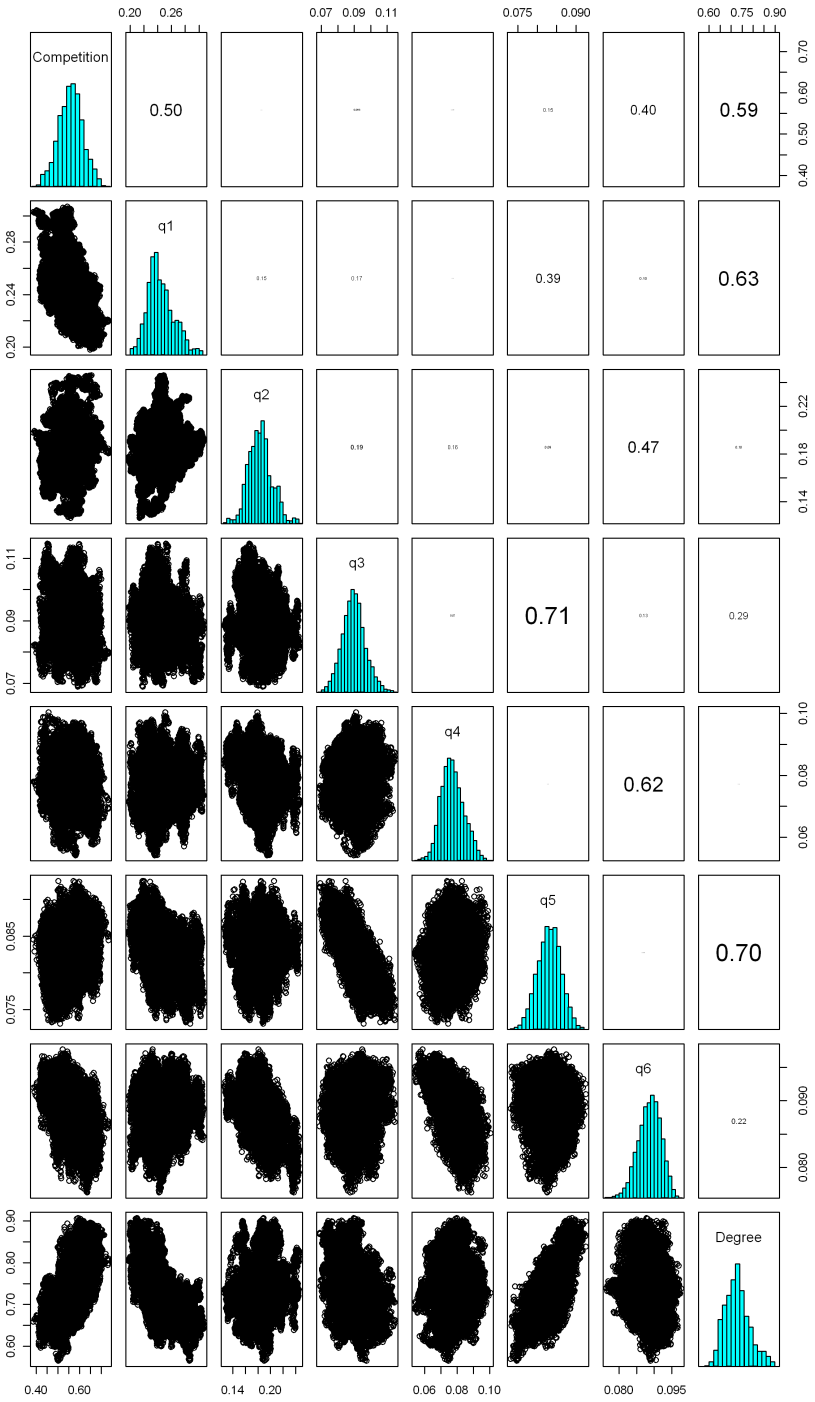


Figure 5.4: Scenario 1: Correlation of the model parameters. Histograms of the posterior distributions are shown on the diagonal. The lower panel holds scatter plots of the two respective variables and the upper panel holds pearsons correlation coefficient. Data is taken after a burn-in period of 10,000.

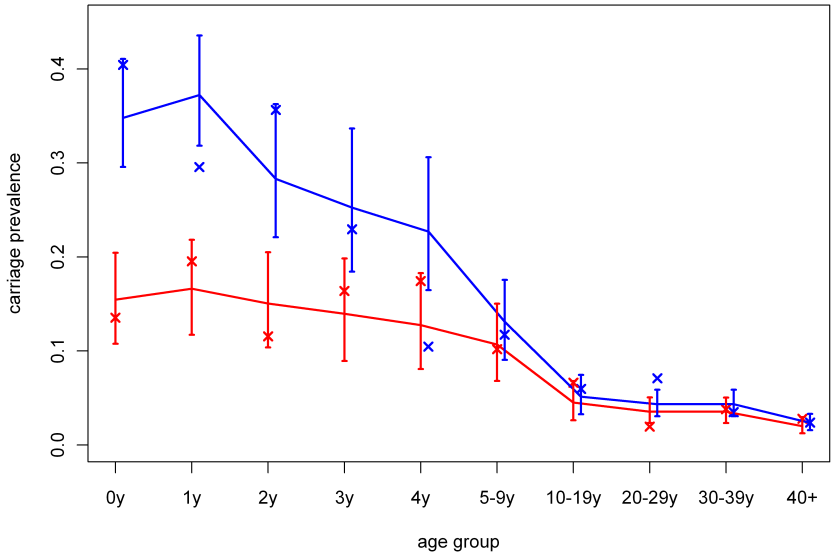


Figure 5.5: Scenario 1: Comparison of observed pre-vaccination VT and NVT carriage prevalence from an English study [6] (blue and red crosses) and the model fit represented by the respectively coloured line and the 95% credible intervals. Age groups are: 0y, 1y, 2y, 3y, 4y, 5-9y, 10-19y, 20-29y, 30-39y and 40+.

Table 5.1: Point estimates and 95% credible interval for the reduction in acquisition due to VT carriage, the age dependent risk of an infectious contact with either VT and NVT and the Vaccine effectiveness against carriage for the three scenarios of interpretation of post vaccination IPD.

	Scenario 1 (trend adjusted) (type 1 excluded)	Scenario 2 (not trend adjusted) (type 1 included)	Scenario 3 (trend adjusted) (type 1 included)
1-c	44% [35,55]	26% [16,36]	0.6% [0.0, 3.9]
q1	0.243 [0.214, 0.275]	0.231 [0.207, 0.261]	0.224 [0.193, 0.252]
q2	0.090 [0.076, 0.104]	0.089 [0.077, 0.103]	0.166 [0.144, 0.198]
q3	0.083 [0.077, 0.089]	0.081 [0.076, 0.086]	0.082 [0.071, 0.093]
q4	0.182 [0.147, 0.213]	0.174 [0.149, 0.201]	0.073 [0.060, 0.086]
q5	0.076 [0.063, 0.092]	0.075 [0.062, 0.089]	0.080 [0.074, 0.086]
q6	0.089 [0.084, 0.095]	0.087 [0.081, 0.093]	0.084 [0.077, 0.089]
VEc	72.3% [62.6, 82.7]	67.9% [59.7, 76.0]	75.8% [66.4, 87.5]

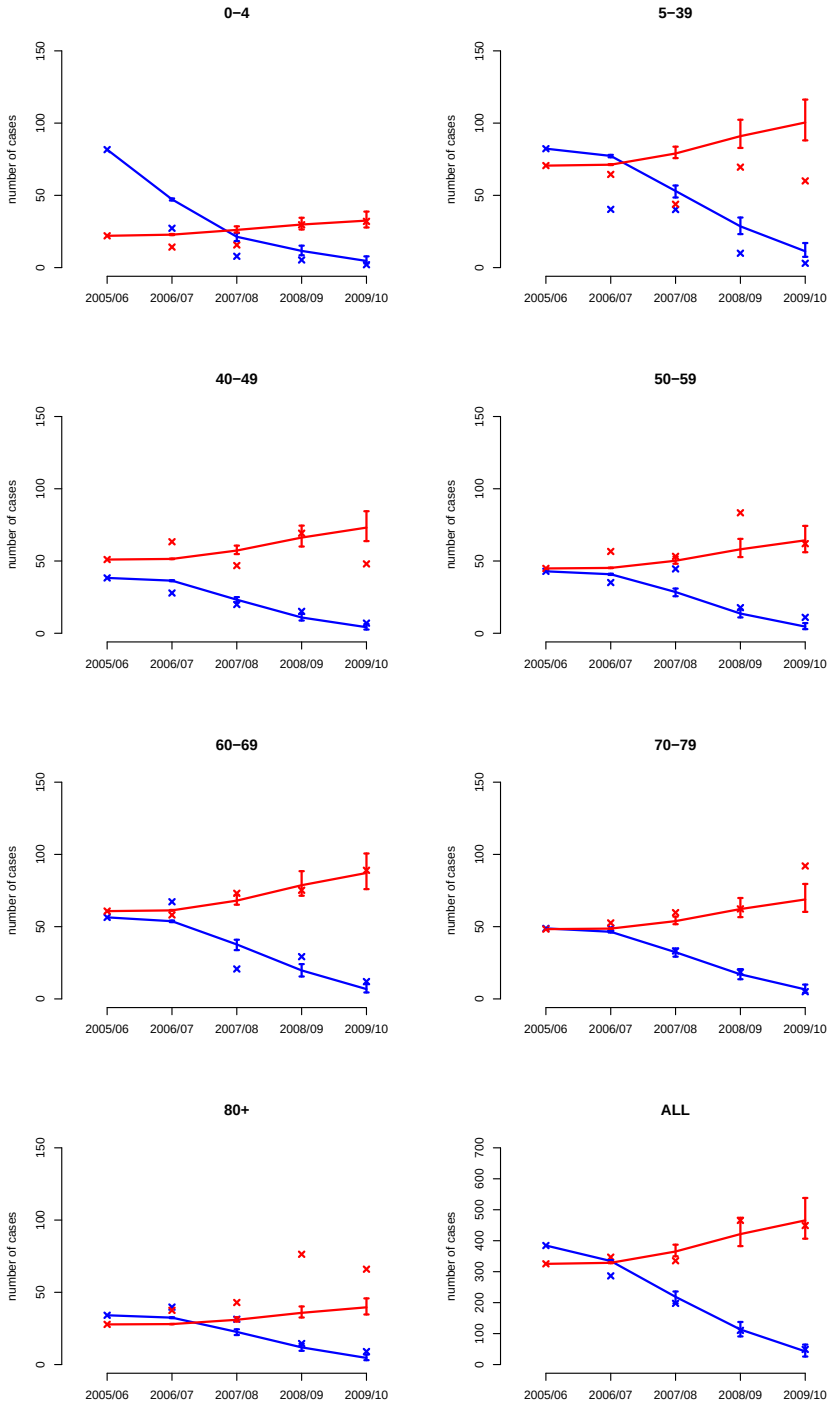


Figure 5.6: Scenario 1: Comparison of observed number of VT and NVT IPD cases (blue and red crosses) and the model fit represented by the respectively coloured line including 95% credible intervals.

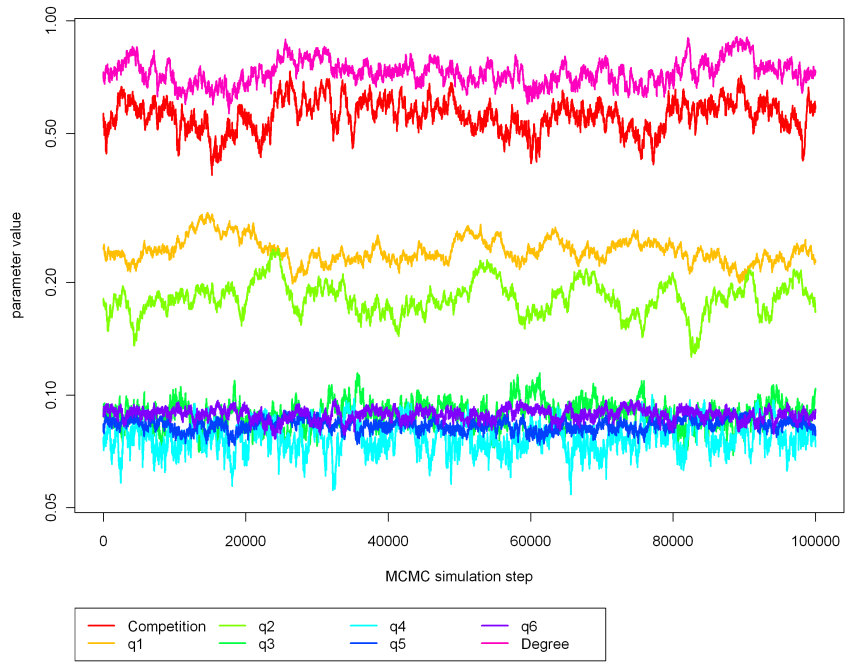


Figure 5.7: Scenario 2 (not adjusted for trends; inclusion of ST1): Markov Chains for fitting the parameters competition, degree and the six q-values.

The reduction of acquisition given colonisation was estimated to be lower than in the scenario 1: 26%, 95% credible interval: [16, 36]. The risk of an infectious contact was rather similar. The model fit in scenario 2 suffered from similar problems as the ones described in scenario 1. However, in scenario 2 the difference between the model and the data is worse in the 5-39 year old and also the increase in NVT IPD in the 0-4 year olds seems not to be captured (see figure 5.10).

Including serotype 1 in the analysis and adjusting for trends in the pre-vaccination era (scenario 3) led to slightly different parameter estimates (see figures 5.11 and 5.12). The carriage data is fitted well (see figure 5.13). Since the data if interpreted this way don't show any evidence of serotype replacement, the reduction of acquisition given colonisation was estimated to be practically zero (with a bi-modal distribution) : 0.6%, 95% credible interval: [0.0, 3.9]. The risks of an infectious contact were again similar to the other scenarios. The risks for an infectious contact with VT in 0-1y, 2-4y and 5+ were 0.224 [0.193, 0.252], 0.166 [0.144, 0.198] and 0.082 [0.071, 0.093]. The risks for an infectious contact with NVT were 0.073 [0.060, 0.086], 0.080 [0.074, 0.086] and 0.084 [0.077, 0.089] respectively. The estimate for the protective effect of PCV7 against carriage was: 75.8% [66.4, 87.5]. The model fit in scenario 3 suffered from a marked decrease in the number of NVT cases after 2005/06 especially in the age group 5-39 which

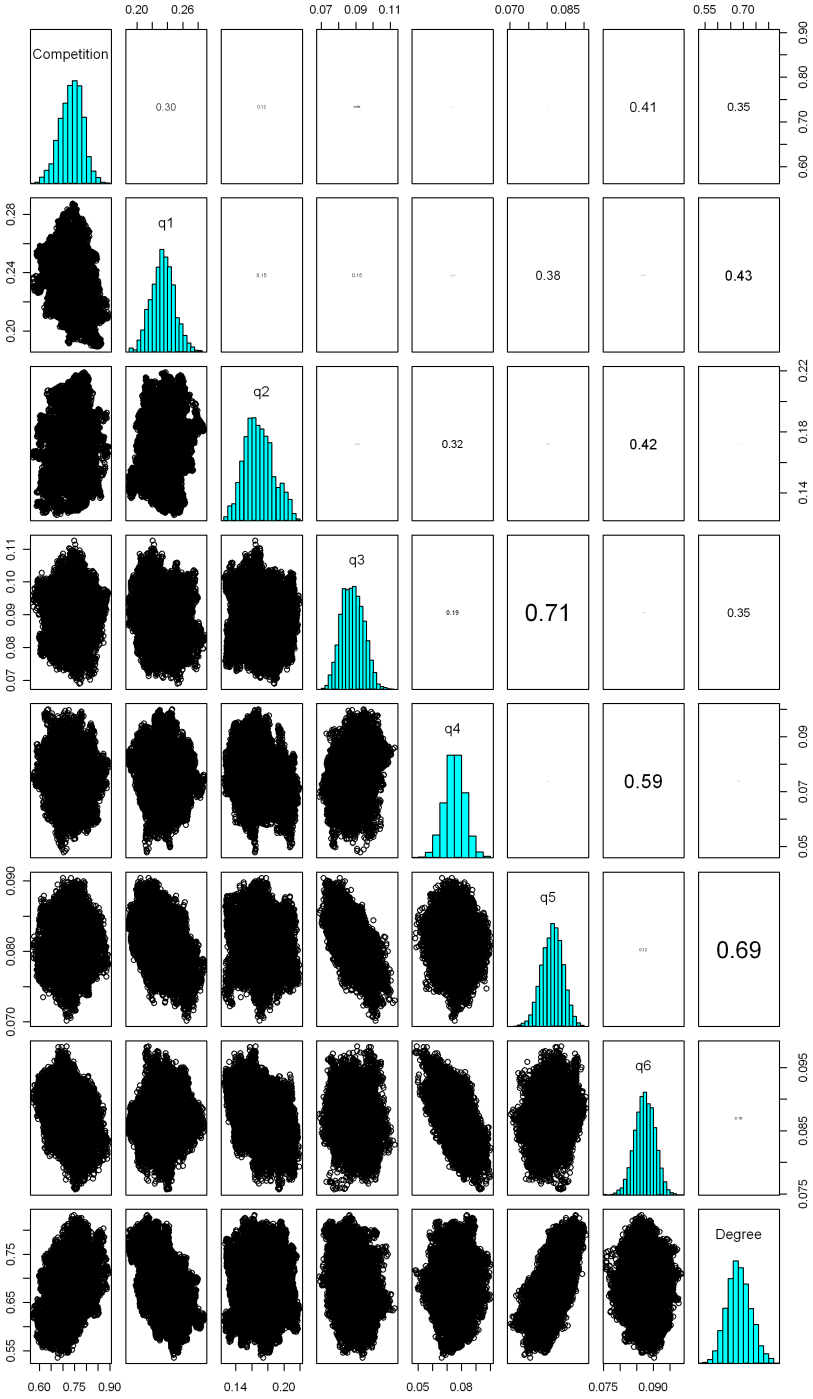


Figure 5.8: Scenario 2: Correlation of the model parameters. Histograms are shown on the diagonal. The lower panel holds scatter plots of the two respective variables and the upper panel holds pearsons correlation coefficient. Data is taken after a burn-in period of 10,000.

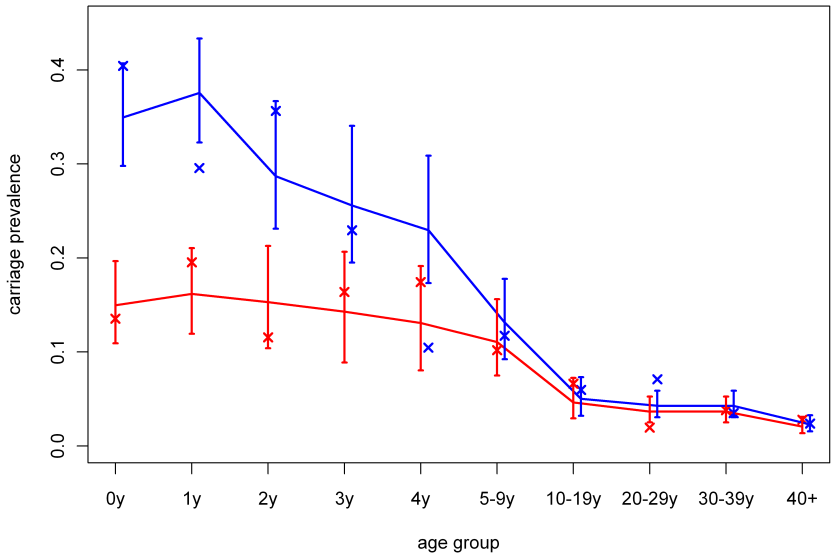


Figure 5.9: Scenario 2: Comparison of observed pre-vaccination VT and NVT carriage prevalence from an English study [6] (blue and red crosses) and the model fit represented by the respectively coloured line and the 95% credible intervals. Age groups are: 0y, 1y, 2y, 3y, 4y, 5-9y, 10-19y, 20-29y, 30-39y and 40+.

is mainly due to the epidemiology of serotype 1 (see figure 5.14).

5.2.3 Discussion

Here an adaptation of the work of Choi et al. [2] to Scottish data on invasive disease in the four years following the introduction of PCV7 is presented. The fitting procedure is changed into a Markov Process offering the opportunity to obtain credible intervals in addition to point estimates for all model parameters. Since the post vaccination epidemiology in Scotland is difficult to interpret this was done by incorporating three scenarios of possible post vaccination epidemiology (compare section 3.1.2). The model fitted well to the post vaccination epidemiology observed in VT disease but failed to mirror the differences in NVT disease replacement observed amongst the age groups.

Convergence of the Markov chains was found to be independent of the starting values. Some of the estimated parameters were found to be correlated, however while a re-parameterisation of the model might yield better mixing or faster convergence of the Markov chain this should not affect the parameter estimates

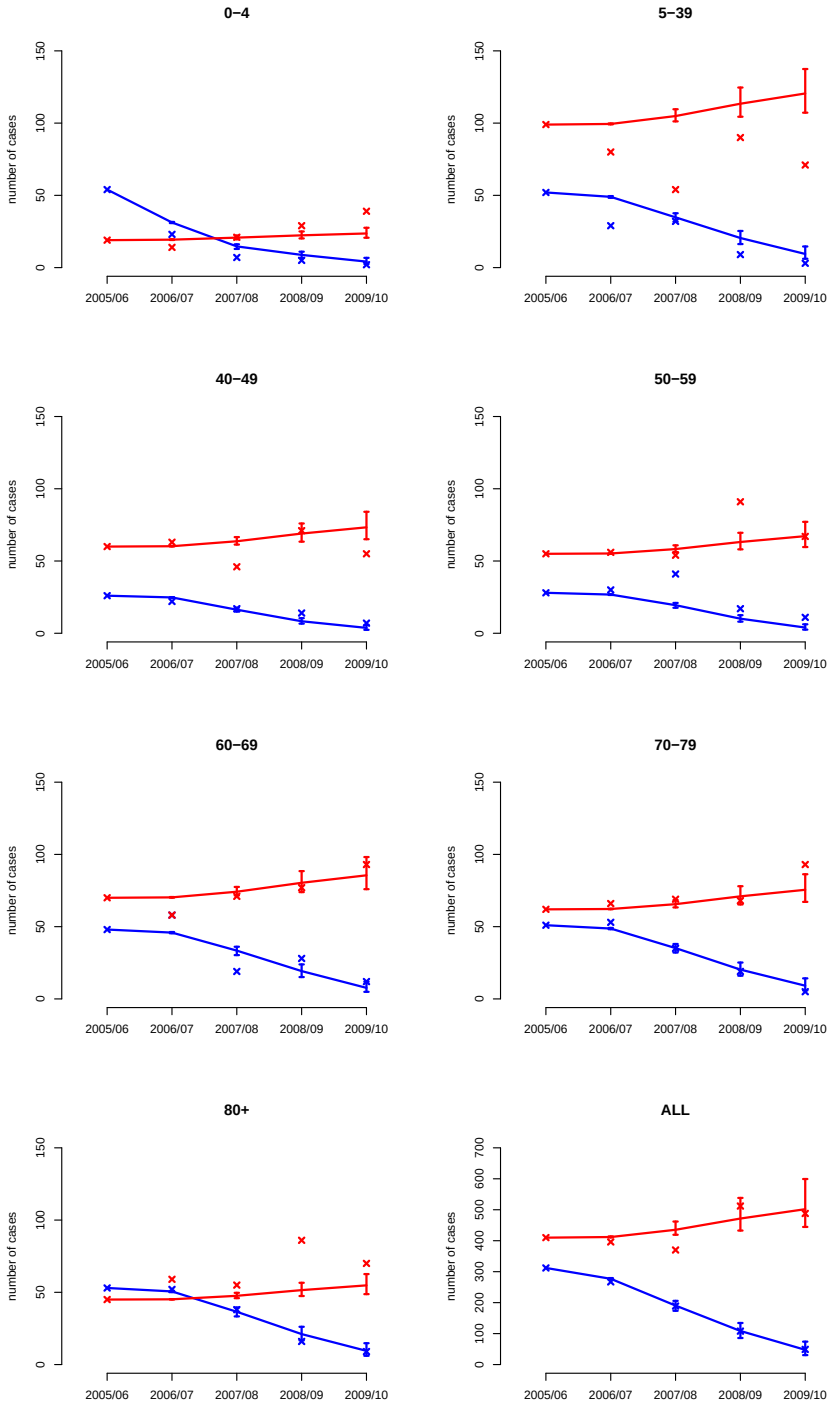


Figure 5.10: Scenario 2: Comparison of observed number of VT and NVT IPD cases (blue and red crosses) and the model fit represented by the respectively coloured line including 95% credible intervals.

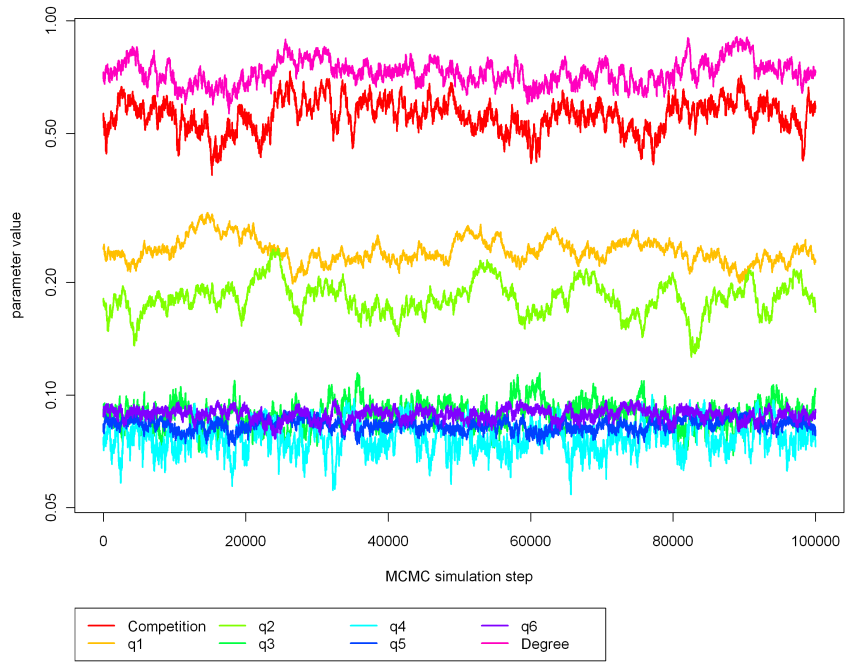


Figure 5.11: Scenario 3 (trend-adjusted; inclusion of ST1): Markov Chains for fitting the parameters competition, degree and the six q-values.

themselves. Not much effort was put into fine tuning of the Markov Chains here to enable faster mixing and convergence, however, the limited complexity of the fitting problem enabled a brute force approach. 100,000 Markov Chain simulations took an average of 3-4 days on one core of a Xeon 5190 (3GHz).

The data on invasive pneumococcal disease in Scotland which is used here for model calibration is not straight forward to interpret; trends prior to vaccine introduction as well as cyclic behaviour of serotype 1 could have influenced the post-PCV7 IPD data and therefore introduce uncertainty in the vaccine induced effects (compare chapter 3.1.2). Presented here are three scenarios, one adjusting for trends and excluding ST1, one including ST1 and not adjusting for trends and one adjusting for trends and including ST1. Competition in these scenarios was estimated to descend by scenario number; in scenario 3 essentially no competition was found, whereas in scenario 1 where an approximate 50% reduction in acquisition of additional colonisation while infected was estimated. However even this is low when compared to other estimates [2, 12]. Compared to the analysis on unadjusted data in scenario 2, scenario 3 shows less competition induced by less replacement when adjusting for pre vaccination trends (see section 3.1.1). Compared to scenario 3, scenario 1 finds increased replacement through the exclusion of ST1 IPD which decreased during the post vaccination era and masked

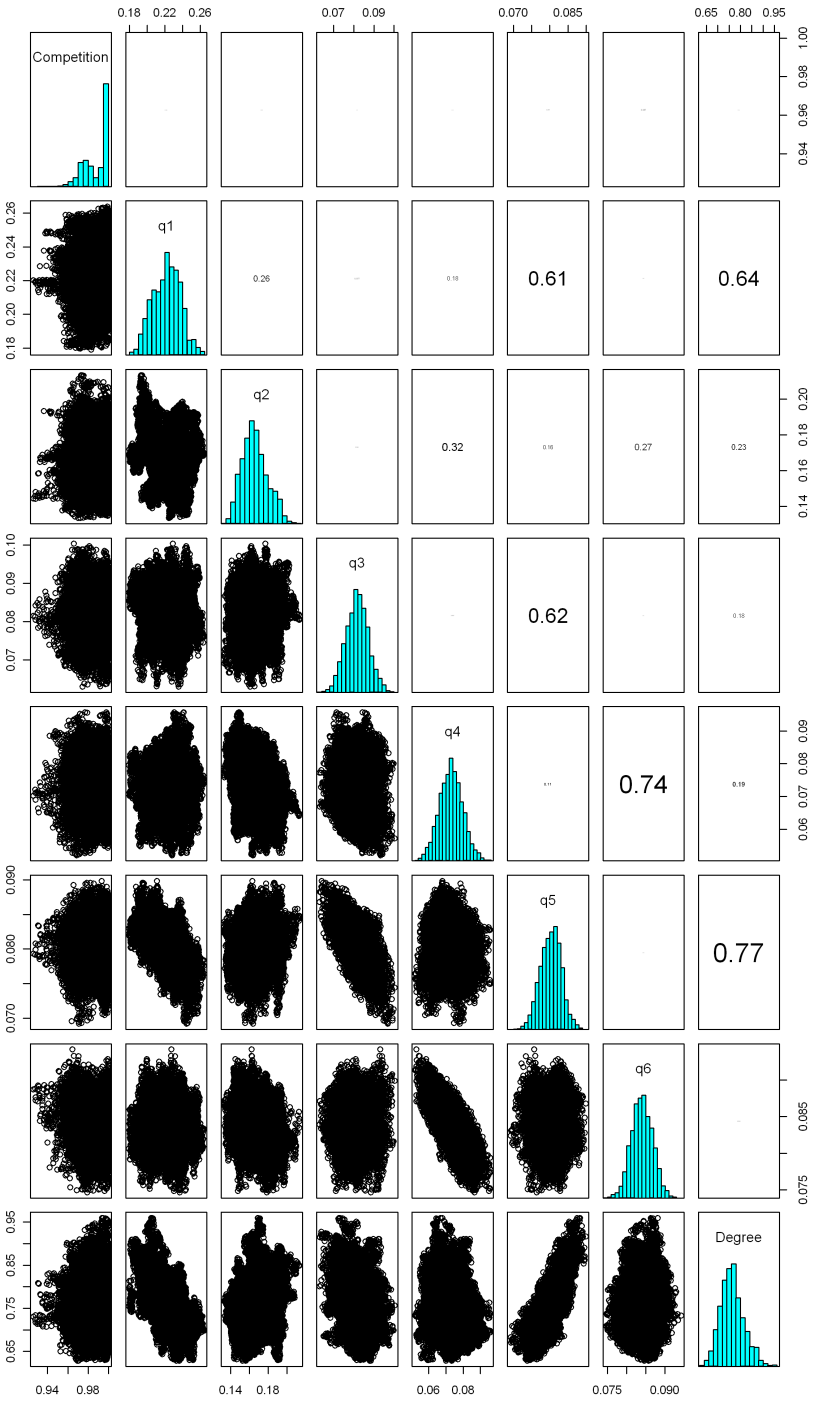


Figure 5.12: Scenario 3: Correlation of the model parameters. Histograms are shown on the diagonal. The lower panel holds scatter plots of the two respective variables and the upper panel holds pearsons correlation coefficient. Data is taken after a burn-in period of 10,000.

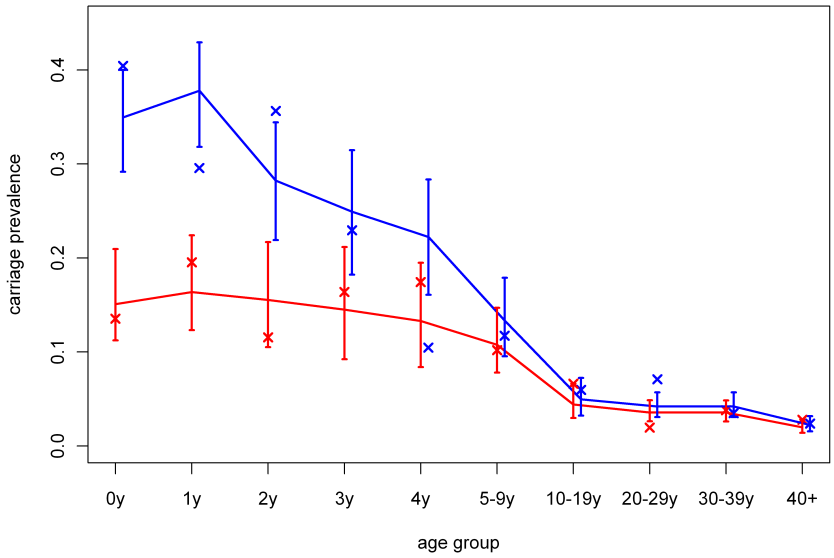


Figure 5.13: Scenario 3: Comparison of observed pre-vaccination VT and NVT carriage prevalence from an English study [6] (blue and red crosses) and the model fit represented by the respectively coloured line and the 95% credible intervals. Age groups are: 0y, 1y, 2y, 3y, 4y, 5-9y, 10-19y, 20-29y, 30-39y and 40+.

any other NVT increase. No formal comparison of the goodness of fit of the model to the three scenarios was undertaken because none of them has captured an important feature of the post PCV7 IPD epidemiology in Scotland, the strong replacement effect in the elderly population with little evidence for replacement amongst other age groups.

Although generally in good agreement with the data for the overall population in all three scenarios, the model showed poor fitting in some age groups. The model is set up with age non-specific competition parameters and assumes that the invasiveness of the NVT group is the same pre and post vaccination. Therefore any replacement in the model is driven by a decrease of VT carriage in vaccinated children and a concurrent increase of NVT carriage which then spreads to the rest of the population. However, in the Scottish data there is evidence for only moderate replacement in children and even less in the 5-39 year olds with almost full replacement observed in the elderly population. This behaviour could possibly be captured with a model by introducing age dependent competition parameters, which in return would allow for age differences in magnitude of replacement effects. However, there is no evidence supporting that protection

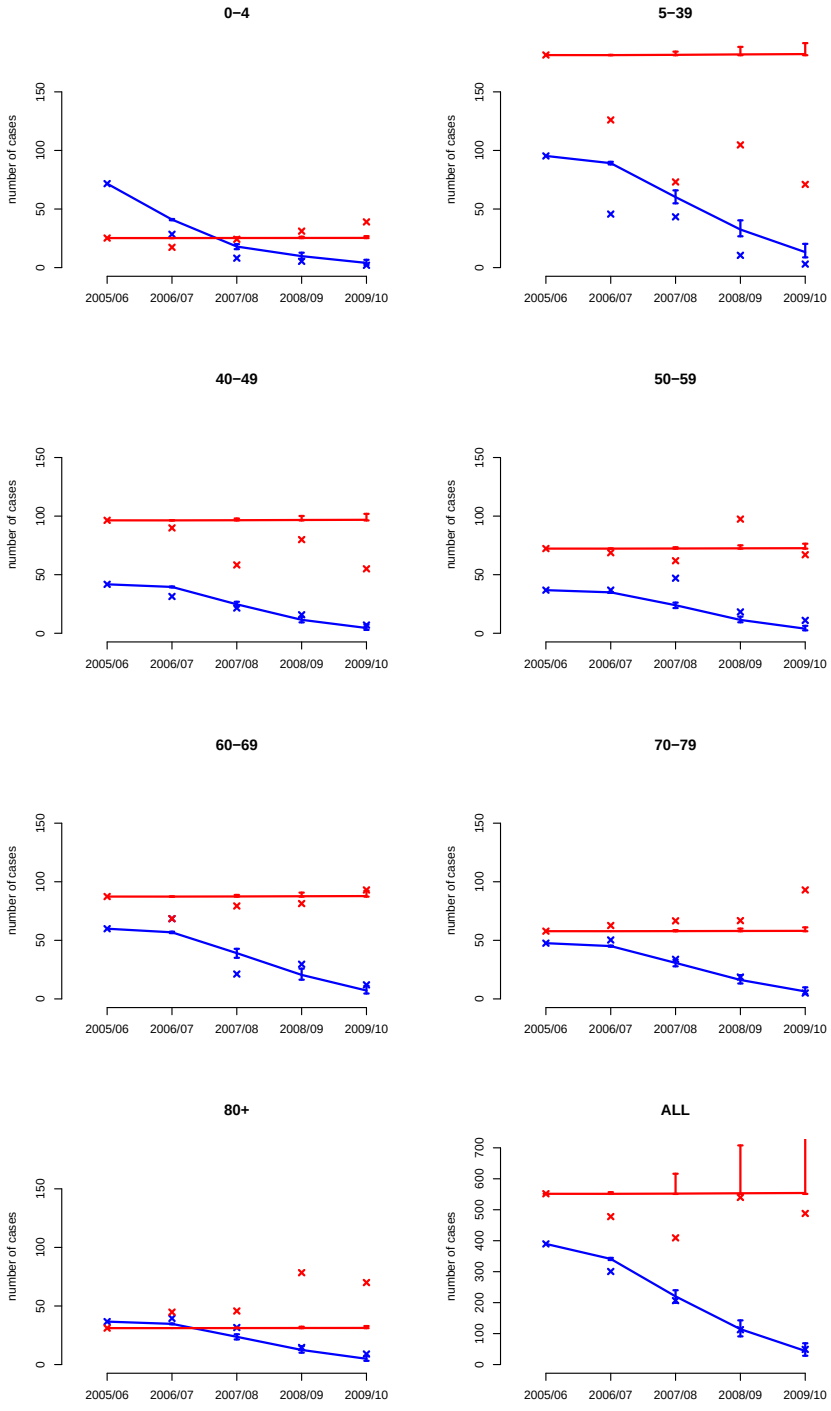


Figure 5.14: Scenario 3: Comparison of observed number of VT and NVT IPD cases (blue and red crosses) and the model fit represented by the respectively coloured line including 95% credible intervals.

against acquisition induced by vaccine type carriage is age dependent. Whether the failure of the model to fit the age-specific IPD data is a result of an insufficiently flexible model parameterisation, an inapplicable mode structure or a deficient understanding of the underlying changes in nasopharyngeal colonisation which led to these observations in IPD (the composition of the NVT groups in some age groups and therefore its invasiveness could have changed following vaccination - compare section 3.1.2) is to be determined.

The model fitted to post PCV7 IPD data from the US was shown to fit the age stratified IPD data closely [1]. No such age stratification was published for its adaptation to fit post PCV7 IPD data from England and Wales [2]. However, the analysis on the impact of the 7-valent conjugate vaccine in E&W showed that replacement effects were highest in children and slightly lower in all other age groups [13]. Therefore the model is likely to be able to capture age stratified post vaccination epidemiology amongst IPD for E&W.

In this work a model previously employed for prediction of the long-term population impact of PCV7 in England and Wales [2] was adapted to Scottish data on invasive pneumococcal disease and a different fitting procedure was employed. The lack of ability of the model to fit the age-specific patterns of NVT replacement in IPD in Scotland highlights the need for a more thorough understanding of the mechanisms driving these observations and yields that for a parameterisation to Scottish IPD data this model is unsuitable in its current form to predict the long-term impact of PCV7. The model used reflects the current knowledge of pneumococcal transmission and disease development although in a simplified version. Its failure to reproduce the observations could come from multiple sources including the failure of the model to capture an essential part of the transmission of the pneumococcus or from an insufficient understanding of the underlying changes is pneumococcal carriage which induced the observed changes in IPD. Other than stochastic effect arising from small numbers, one of the possible reasons to explain the rather unanticipated post vaccination IPD epidemiology of replacement affecting predominantly the elderly population could be a changing NVT serotype distribution, and hence an altered overall invasiveness, in the post-PCV introduction era amongst some age groups. In chapter 3.1.2 it is shown, that there are particular serotypes which are responsible for the majority of replacement and therefore became increasingly prevalent within the group of NVT. In the model it was assumed, that the invasiveness of the NVT group did not change which might be wrong if the major replacing serotypes were more or less

invasive than an average NVT in the pre-vaccination era. Further studies, in particular information on nasopharyngeal carriage could prove useful, and are needed to further evaluate the reasons for this discrepancy.

5.3 Transformation to an individual-based model

PCV13 replaced PCV7 in the UK in April 2010. Prior to its introduction Choi and colleagues expanded the previously implemented model (see figure 5.1) to provide information about the differential benefits of introducing PCV13 over discontinuing pneumococcal conjugate vaccination [14]. Amongst many other changes, including the structure of the model to incorporate the compartments NVT, VT1 (serotypes in PCV7), VT2(serotypes in PCV13 and not in PCV7), the project faced the challenge of representing the vaccine uptake for a mixed schedule of PCV7 and PCV13 given to children at the start of the switch to PCV13. That is that children could have received e.g. a first dose of PCV7 and a continued schedule with PCV13 which would provide them with part protection against the additional serotypes included in PCV13 and full protection against PCV7 serotypes. This would vastly increase the complexity of compartments needed for the deterministic model. However, since the model is only fitted to data from the pre-vaccination era a hybrid approach was taken.

A deterministic model was set up and fitted to the available data then the estimated parameters were fed into an individual based framework for the same model to estimate the long-term effects after PCV13 vaccination with a detailed reflection of vaccine uptake. This approach allowed for both efficient fitting with a deterministic model and easy representation of various levels of protection from mixed schedules in the post-PCV13 era through an individual based model. As part of this PhD study the transformation from the deterministic model, explained in detail in sections 5.1 and 5.2, to its individual based equivalent, and its later adaptation to a more complex structure, was undertaken and is summarised in the following.

The following sections discuss the adequate reflection of chance when transforming transition rates into transmission probabilities, the role of assumptions on age distribution in a stochastic transmission model and the computationally efficient implementation of a simple birth and death process. Furthermore, the relation of the usual expression of the force of infection in a stochastic model and its deter-

ministic version is explored. Finally the results of the analysis of Choi et al. [14] are summarised.

5.3.1 Random numbers

In an individual based model a change of state is determined probabilistically (via random numbers) rather than by fixed rates as in deterministic models. This raises the question about how chance is appropriately reflected in this context.

In day-to-day life observations with underlying mechanisms too complex to understand or predict are considered to happen randomly, although in fact they do not. For example given the initial conditions the outcome of a coin flip could theoretically be predicted by the rotation speed and the distance travelled. However, this can be complicated by what in chaos theory is referred to as the butterfly effect - the sensitive dependence of the outcome on small changes in the initial conditions. The Oxford Dictionary defines (statistical) randomness as "governed by or involving equal chances for each item" where chance is defined by "the occurrence of events in the absence of any obvious intention or cause" [15]. Since any computer program does nothing but progressing signals ("cause") by definition nothing like a truly random number exists in the digital world. The way around this are so called pseudo random number generators. Given a seed (usually set by the current system time) these algorithms generate a sequence of numbers that approximates the characteristics of random numbers. Different generators have been established varying hugely by their approximation quality, sequence length and computational costs. Common classes of these algorithms include linear congruential generators, Lagged Fibonacci generators and linear feedback shift registers. When employing the built-in 'rand()' function in C++ to draw the random numbers needed for the transition probabilities it turned out that this function is insufficiently random and therefore led to results deviating from those of the deterministic model. For this work long sequences of random numbers for millions of individuals are needed to be drawn. Hence, the recently developed Linear Feedbacked Shift Register generator "SIMD-orientated Fast Marsenne Twister" (SFMT) (e.g. available from the RandomLib C++ interface or the GNU Scientific Library) was chosen because it provided a good trade off between computational speed and pseudo-randomness [16].

5.3.2 Population model

In essence there are two commonly employed approaches for constructing the population demographics and the induced transmission dynamics for a deterministic infectious disease model:

- a flat population structure in which each age cohort consists of the same number of individuals and where the proportion infected is weighted by the actual population for the calculation of the force of infection (this is used for the model discussed in sections 5.1 and 5.2), or
- a birth and death process determining a population in steady state which has a population structure more similar to the real one.

Most deterministic modelling approaches use either of them. The problem with the second approach is that with the inclusion of a simple demographic process the real population structure can only be approximated to a certain extent which might lead to misjudgement of the importance of some age groups for the transmission dynamics. However, if uncertainty is intended to be considered, as is the case for an individual based model, the latter approach has the advantage of more properly reflecting differences in uncertainty amongst differently large age groups. That is, that uncertainty in the elderly population (80+ years) is increased by their small population size relatively to other age bands, which would not be adequately reflected if employing a flat population structure. Therefore, a mixture of both approaches is pursued here. The model employs a simple birth and death process to approximate the relative size of the population in different age bands and to infer uncertainty estimates from it and also weights the proportion infected for the calculation of the force of infection by the actual population age distribution in 2007 for England and Wales as obtained by the Office for National Statistics (ONS).

The overall size of the population, N, in the individual based model is determined by the number of newborns and the survival rates. The newborns enter the model as naive individuals being 0 weeks old and fully susceptible. Each individual of age a has a yearly probability of survival $s_a = (1 - m_a)$, where m_a are the yearly mortality rates in E&W in 2007 from the ONS. The maximal age in the model was restricted to one hundred years ($s_{100} = 0$). The weekly probability of an

individual of age a to survive then calculates as:

$$P[x_a \leq s_a^{\frac{1}{h}}],$$

where x is a uniform distributed random number and $h = 48$ is the number of weeks per year ($h = 48$ was chosen to comply with the deterministic model which uses this approximation in order to more easily convert between parameter values measured in weeks and data which is stratified by months).

5.3.3 Population steady state

Before the disease is introduced to the system the population needs to be in a steady state in order to make sure that measured alterations in disease outcome are not induced by changes in the population. A classical approach to reach a population steady state is to just run the model of the birth and death process until time t^* where the overall population $N(t)$ fluctuates around a constant value. This usually introduces a high burden of computational cost to the model before any simulation of infectious disease spread is even introduced. To reduce these costs the population steady state could be computed, but based on a smaller population size, and then used for initialisation of the full population model. Since the variance of the steady state employing the smaller population size will significantly exceed the variance of the full population, while the mean is the same, an average over a sufficient period should be taken for initialisation. The model considered here has a simple population structure; namely a birth and death process with a constant birth rate. Therefore the population steady state can be derived by hand. Let $w \in \{1, \ldots, 48 * 100\}$ be the age of an individual in weeks. The expected number of individuals at age w^* of a birth-death process with b weekly births and a chance of surviving week w_i given by s_{w_i} is calculated as

$$N_{w^*} = b \prod_{i=1}^{w^*} s_i.$$

Since usually $N_{w^*} \notin \mathbb{N}$ assignment of individuals to their age groups in an individual based model according to a calculated population steady state is not straight forward. To overcome this issue N_{w^*} was rounded up or down according to a draw from uniformly distributed random numbers in $[0, 1]$ being smaller than the decimals of N_{w^*}.

5.3.4 Transmission process

The transmission process in a model for the spread of infectious diseases is determined by each individuals' probability of becoming infected in a certain time step - the force of infection (FOI, λ). Following the rationale proposed by Reed and Frost [17] this is equivalent to 1 minus the probability to evade all infectious contacts. The probability to evade all infectious contact then can be interpreted as the probability to evade an infectious contact with each of the infected individuals. Let I be the current number of infectious individuals and β the probability to get infected if you get in contact with one infectious individual. Assuming the probability of getting infected differs neither by the infector nor the infectee the FOI is:

$$\lambda(\beta, I) = 1 - (1 - \beta)^I.$$

Here $\beta \in \mathbb{R}$ and $I \in \mathbb{N}$. For convenience it is assumed in the following that $I \in \mathbb{R}$. The underlying small world assumption (compare "random mixing" in section 1.2) is assumed to hold in order to comply with the deterministic modelling approach. Now $\lambda : (\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ is infinitely differentiable so we can consider Taylor's multi dimensional second order approximation of the Reed Frost equation:

$$\begin{aligned}\lambda(\beta, I) =& \lambda(\beta_0, I_0) \\ &+ \lambda_\beta(\beta_0, I_0)(\beta - \beta_0) + \lambda_I(\beta_0, I_0)(I - I_0) \\ &+ \lambda_{\beta\beta}(\beta_0, I_0)(\beta - \beta_0)^2 + \lambda_{\beta I}(\beta_0, I_0)(\beta - \beta_0)(I - I_0) \\ &+ \lambda_{II}(\beta_0, I_0)(I - I_0)^2 + e(\beta, I)\end{aligned}$$

with λ_x and λ_{xx} being the first and second order derivative of λ in x and $e(\beta, I)$ a third order error term. The corresponding derivatives calculate as:

$$\begin{aligned}\lambda_\beta(\beta, I) =& I(1 - \beta)^{I-1} \\ \lambda_I(\beta, I) =& ln(1 - \beta)(1 - \beta)^I \\ \lambda_{\beta\beta}(\beta, I) =& I(I - 1)(1 - \beta)^{I-2} \\ \lambda_{\beta I}(\beta, I) =& 1(1 - \beta)^{I-1} + I\ ln(1 - \beta)(1 - \beta)^{I-1} \\ \lambda_{II}(\beta, I) =& ln(1 - \beta) ln(1 - \beta)(1 - \beta)^I.\end{aligned}$$

At $(\beta_0, I_0) = (0, 0)$ all first and second order derivatives except $\lambda_{\beta I}(\beta_0, I_0) = 1$ are zero, so

$$\begin{aligned}\lambda(\beta, I) &\approx 0 + 0 + 0 + 0 + 1(\beta - 0)(I - 0) + 0 \\ &= \beta I.\end{aligned}$$

This gives a reasonably good and more importantly computationally less costly and numerically more stable approximation of the force of infection for a neighbourhood around $(0, 0)$. In fact this approximation is the way the force of infection is represented in deterministic models (compare section 1.2). It then interprets as the proportion of susceptibles becoming infected.

Both approaches can be refined for different transmission patterns for several subgroups. Consider the population to be divided into n subgroups (in this case these are age groups from the Polymod mixing patterns). In the deterministic (Kermack & McKendric) approach the subgroup specific FOI calculates as:

$$\lambda_i = \sum_{j=1}^{n} \beta_{j,i} I_j \ , i \in \{1, \ldots, n\},$$

where $\beta_{j,i}$ denotes the probability of an effective contact of an individual of subgroup j to one of group i and I_j denotes the number of infective people in group j. In the Reed Frost approach this translates to :

$$\lambda_i = 1 - \prod_{j=1}^{n} (1 - \beta_{j,i})^{I_j} \ , i \in \{1, \ldots, n\}.$$

As the aim of this individual based modelling approach was to reproduce the results of a given deterministic model anyway, the second order approximation of the FOI about $(0, 0)$ was used to determine whether each individual was infected. However, validation against the actual Reed Frost approach could not detect any differences between these approaches for the model in question.

5.3.5 Assessing the potential impact of PCV13 in E&W

In all remaining aspects, the individual based model was built to follow exactly the deterministic approach. When parametrised according to the deterministic model for PCV7 the individual based model did give consistent results with only

minor impact of stochastic effects provided that the model population exceeded a threshold of approximately 100,000 individuals. The following paragraphs provide a short summary of the methods and findings of the work of Choi et al. [14] where this individuals based model was partly used for.

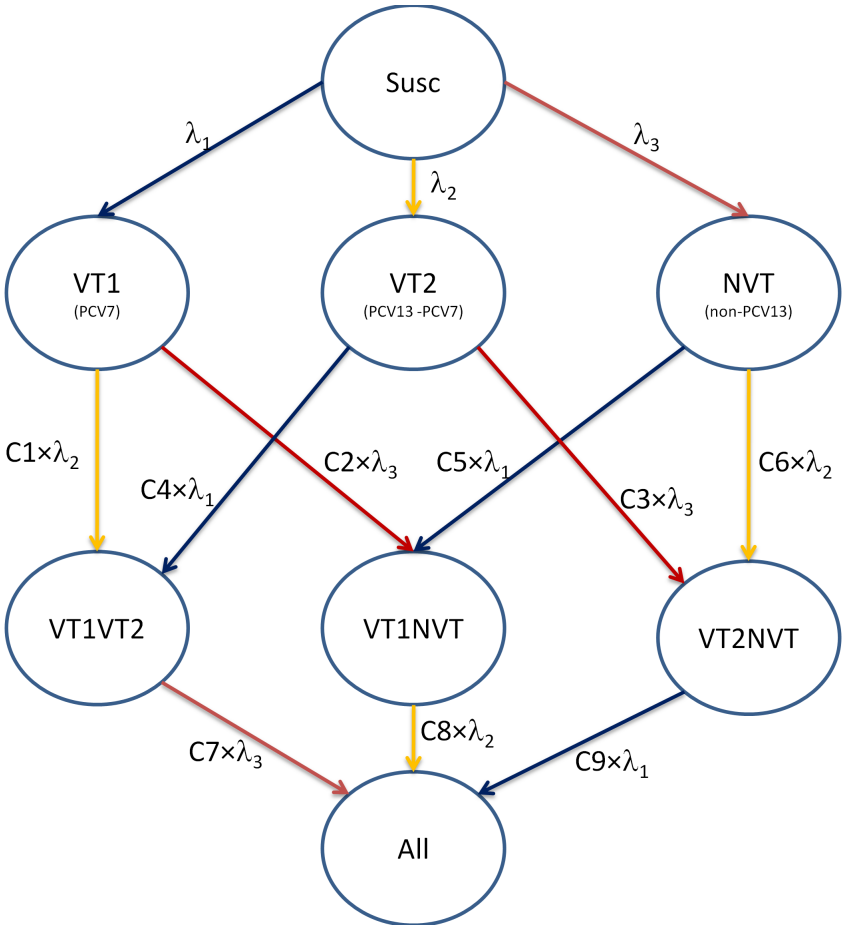


Figure 5.15: Flow diagram showing the extended number of infection states over the models discussed in sections 5.1 and 5.2 to incorporate vaccination with PCV13 replacing PCV7. VT1VT2, VT1NVT and VT2NVT are the states of dual colonisation and "All" refers to infection with VT1, VT2 and NVT at the same time. Parameter naming conventions follow these from figure 5.1 (from Choi et al. [14]).

The individual based model was adapted to comply with the expanded deterministic model representing the transmission dynamics of the group of serotypes included in PCV7 (VT1), the additional types targeted by PCV13 (VT2) and the remaining types (NVT) [14]. Figure 5.15 (compare to Figure 5.1) provides an overview of the different carrier states considered (being infected with no, one, two or three serotypes of the different groups at a time), their interaction and the naming conventions for the forces of infection $\lambda_1, \lambda_2, \lambda_3$ and the competition parameters $C_1, \ldots, C_9$. Data on vaccination uptake was employed as before and

following introduction of PCV13 was assumed to stay at levels similar to those observed 3 years after introduction of PCV7. Individuals can be unprotected, partially protected or fully protected against the PCV7 serotype group in the deterministic model and unprotected, partially protected or fully protected independently against the VT1 and VT2 groups in the individual based model. Individuals receiving a single dose above the age of 1 year as a catch up were considered fully protected. In the regular schedule children receiving their first dose under the age of 1 year were considered partially protected, any additional dosage then had them fully protected. In individuals on a mixed schedule of PCV7 and PCV13 similar things applied; the first dose grants partial protection and each additional dose full protection against the respective serotype groups. E.g. an individual who has received a single dose of PCV7 around April 2010 and a dose of PCV13 2 months later was considered fully protected against VT1 but only partial protected against VT2. Protection was considered to be subject to general waning at a rate reflecting duration of full vaccine protection of 10 years. Vaccine effectiveness against carriage of either of the two vaccine serotype groups was assumed to be 52% for fully protected individuals and 26% for partial protected individuals. Protection against progression to disease was assumed to be 100%. The complexity of the different vaccination histories made this hybrid approach useful, since it would have introduced many additional compartments to the deterministic model and could easily be reflected in the individual based model.

The deterministic model was run for parameter estimation as presented in Choi et al. [14] in a very similar fashion to the earlier model [2]. In brief: The Poisson deviance of the pre-vaccination model to pneumococcal carriage data from England [6] was calculated. Case:carrier ratios for the three serotype groups and 16 age groups were derived from the modelled carriage incidence and data on IPD in 2005/06. Poisson deviances of the number of predicted model cases and IPD data from the three years following implementation of PCV7 were calculated. A set of optimal parameters were derived by minimising the overall Poisson deviance through a simplex algorithm. The parameter estimates were then transferred to the corresponding individual based model. Starting at the date of PCV13 vaccination (April 2010) the individual based model was used to simulate 25 years of post PCV13 carriage and used the estimated case:carrier ratios to infer IPD.

Two aspects of uncertainty were included for model predictions, (i) the adjustment for pre vaccination trends and (ii) the amount of competition introduced

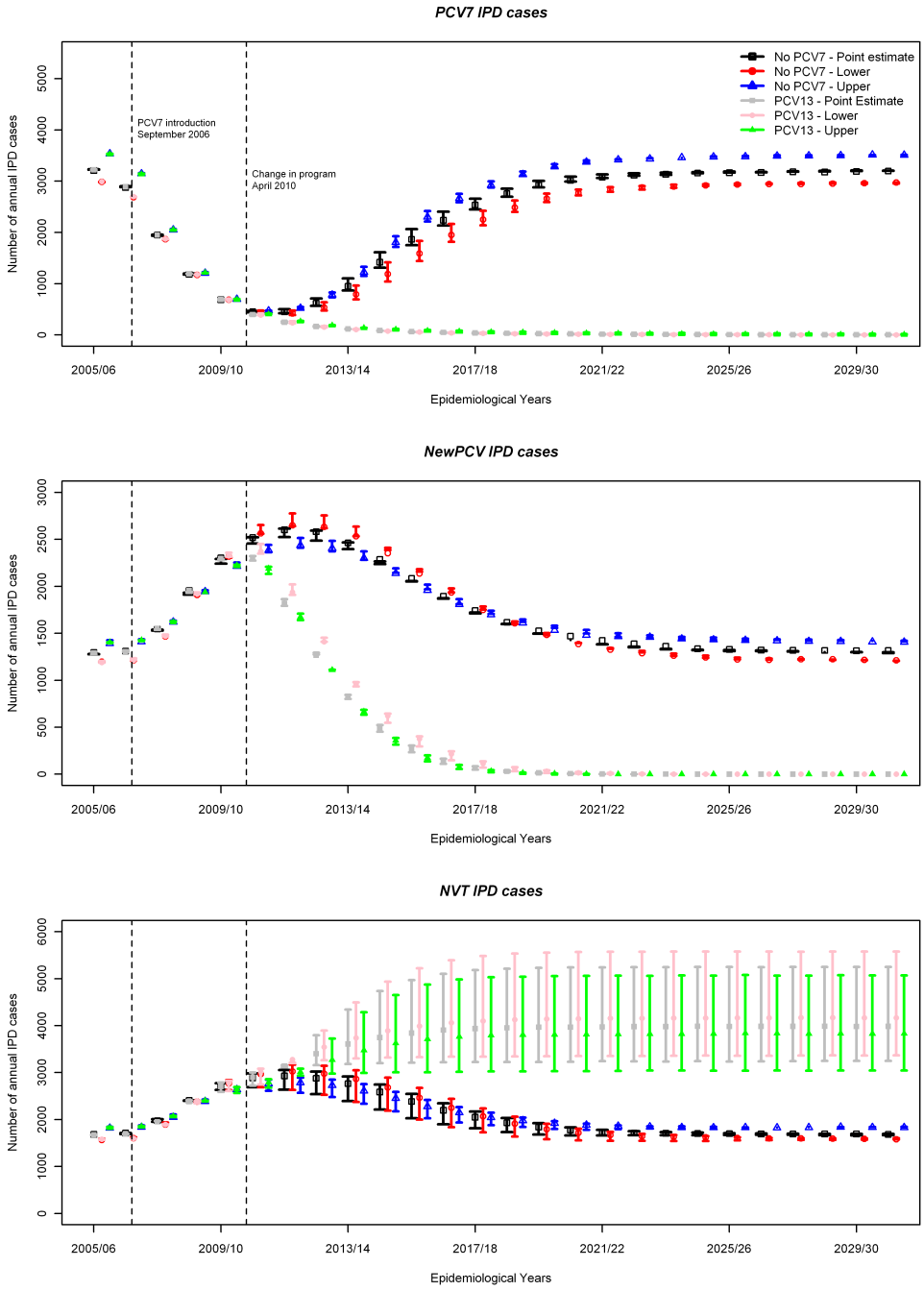


Figure 5.16: Model projections of the long-term effects of either stopping pneumococcal conjugate vaccination or replacing PCV7 with PCV13 on April 2010 on the number of IPD cases caused by VT1, VT2, and NVT serotype groups in England and Wales over 35 years since the introduction of PCV7. C3 (representing the protective effect of carrying VT2 serotypes against infection with NVT) takes the values 0, 0.5, and 1 in different scenarios. Filled circles show results with C3=0.5 and the error bars represent results with C3=0 or 1. Scenarios are presented in terms of IPD data being adjusted by high, low and medium estimates for the secular trend in IPD prior to PCV7 introduction (from Choi et al. [14]).

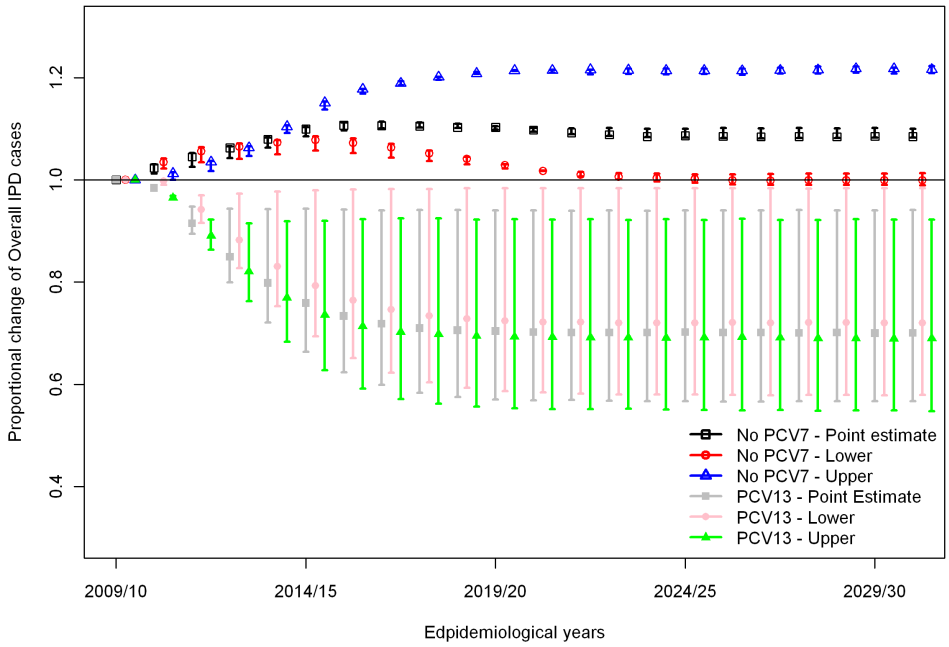


Figure 5.17: Proportional changes in the number of overall IPD cases following either stopping pneumococcal conjugate vaccination or replacing PCV7 with PCV13 on April 2010. PCV7 is replaced with PCV13 six years after PCV7 introduction. C3 (representing the protective effect of carrying PCV13-PCV7 against infection with NVT) takes the values 0, 0.5, and 1 in different scenarios. Filled circles show results with C3=0.5 and the error bars represent results with C3=0 or 1 (from Choi et al. [14]).

on acquisition with NVT by carriage with VT2.

(i) Similarly to the trend adjustment on Scottish IPD data in section 5.2.1 data on IPD in England and Wales were adjusted for pre vaccination trends. The point estimates and the boundary values of the 95% confidence interval on the estimated trend provided three scenarios for IPD data adjustment, referred to as "lower", "point" and "upper". The "lower" scenario hereby corresponds closely to no data adjustment.

(ii) One key aspect determining the overall impact of PCV13 vaccination is the amount of replacement likely to occur as a result of reduced competition induced by carriage from the six additional serotypes in PCV13. In the absence of reliable information about multiple carriage episodes this has been estimated by the inclusion of some years of post vaccination data for PCV7 [1,2]. At the time of the analysis no data about post PCV13 replacement was available which resulted in the insensitivity of the goodness of fit to $C3$ (the competition on NVT induced by the serotype group VT2; see figure 5.15). Therefore the parameter space of $C3$ ($C3 \in [0, 1]$) was explored in three scenarios, $C3 = 0, 0.5, 1$ even though most previous estimations for serotype group competition were found to be closer to competitive exclusion ($C \in [0, 0.5]$).

In figure 5.16 the estimated number of cases of VT1, VT2 and NVT IPD following either vaccination with PCV13 or no vaccination after 2010 are presented. Stopping conjugate vaccination altogether is predicted to lead to the same levels of invasive disease as before the start of conjugate vaccination. Replacing PCV7 with PCV13 is predicted to lead to further decreasing incidence of the types included in PCV7 and to reduce the number of IPD cases associated with the additional six serotypes in PCV13 to essentially zero within eight years after introduction of PCV13. However, non-vaccine serotype IPD is likely to increase following vaccination with PCV13. The extent of this increase is largely dependent on the amount of competition induced by VT2 carriage on NVT carriage acquisition. Despite the uncertainty about the magnitude of serotype replacement the introduction of PCV13 is predicted to further decrease the overall burden of invasive disease in all scenarios, while stopping conjugate vaccination is likely to increase this burden (see figure 5.17). The fact that, even in the scenario of full serotype replacement in carriage (very high competition $C3 = 1$), the overall burden of IPD is predicted to be reduced for all scenarios implies that the invasiveness of

the replacing serotypes is estimated to be lower than that of the VT2. Figure 5.18 shows that this is the case particularly in the elderly population where most of the disease burden is. This adds to the findings from section 3.1.3, where a lower invasiveness of the non PCV13 serotypes was estimated in the population younger than 60 years.

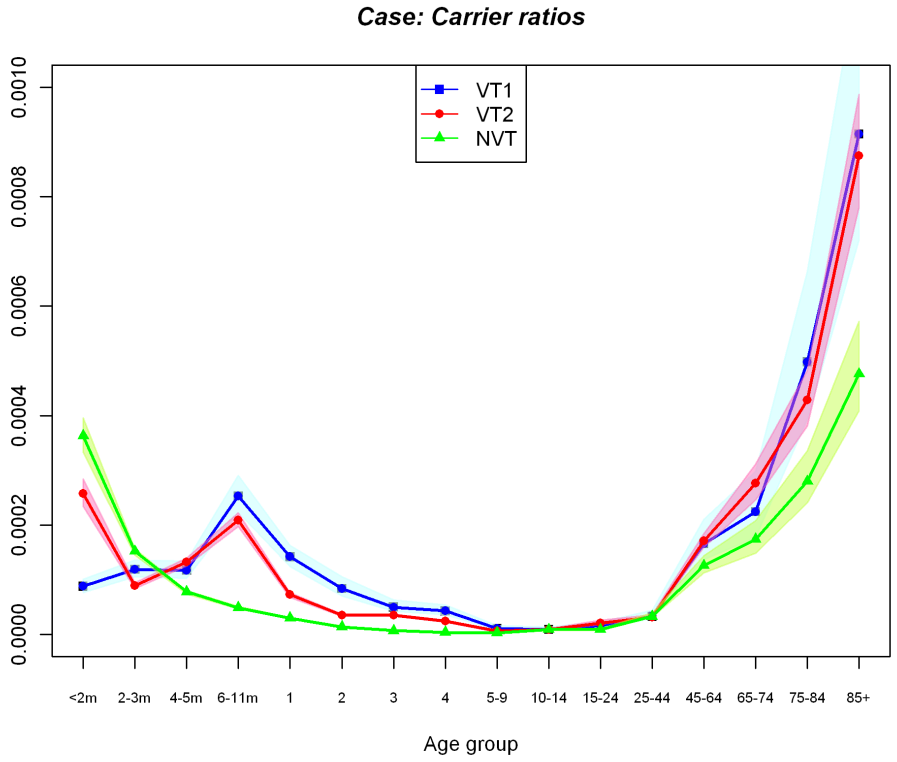


Figure 5.18: Propensity for individuals carrying pneumococcal serotypes in three serogroups (VT1, VT2 and NVT) among 16 age groups to develop disease. The ratios obtained from the scenario using the competition parameter, C3, assumed to b2 0.5 and the IPD cases adjusted using the point estimate of the increasing trend (shaded areas show the ratios estimated using the lower and upper boundaries of its 95% CIs) (from Choi et al. [14]).

Summary

In this section the predicted impact of inclusion of PCV7 into the national immunisation schedule in England and Wales was reviewed. Early estimates based on post vaccination data from the US yielded much less serotype replacement than observed shortly after introduction of PCV7. Using post vaccination IPD data from England and Wales the long term impact was thought to be lower

then estimated previously but still to be beneficial in reducing the burden of IPD. This model was transferred to enable fitting to data on vaccine uptake and invasive pneumococcal disease in Scotland and embedded in a Bayesian parameter estimation framework to allow the assessment of associated uncertainty. Three scenarios were considered: fitting to Scottish IPD data (i) adjusted for pre vaccination trend and excluding disease caused by the volatile serotype 1, (ii) unadjusted and (iii) adjusted for pre vaccination trend but including serotype 1 disease. The fitting process behaves well and reasonable fits to colonisation data and overall IPD are achieved in all scenarios. However, discrepancies between model estimates and age stratified IPD data could not be overcome because of the nature of the model: NVT replacement is thought to happen in the age group which is most important for the transmission dynamics (school age children) and from there spread to other age groups and is reflected as such in the model. This is contradictory to the observation in Scotland of most replacement being observed in the elderly population. The reason for this is unknown (see chapter 3.1.2) but could include differences in the contribution of replacing serotypes to invasive disease by age or stochastic effects in the data. The model was further transferred into an individual based model and caveats in this process were discussed. This model then allowed the flexibility to incorporate a detailed representation of vaccine protection following the switch from PCV7 to PCV13 in April 2010.

These results further highlight the need of a proper understanding for the underlying ecology in order to describe the dynamics of replacement on the biologically plausible basis and make sure that the model in use is an adequate reflection of these. Also the pooling of serotypes into VT and NVT groups introduces limitations to these models. With the pronounced emergence of some NVTs the serotype distribution and hence the invasiveness of the NVT group in some age groups may have changed. This could have lead to the distinct post vaccination epidemiology in Scotland, where serotype replacement in IPD has been mainly confined to the elderly population, and cannot be captured by models pooling all NVTs. Hence, in the following section a model is built which aims to reflect the diversity of pneumococcal serotypes and tries to investigate a possible link between the ecological patterns of competition and coexistence and pneumococcal immunity.

Bibliography

[1] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, et al. (2010) Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC infectious diseases 10: 90.

[2] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[3] Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. (2003) Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. The New England journal of medicine 348: 1737–46.

[4] Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease after pneumococcal vaccination. Lancet 6736: 1–12.

[5] Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine 5: e74.

[6] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[7] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[8] Weinberger DM, Trzccinski K, Lu YJ, Bogaert D, Brandes A, et al. (2009) Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathogens 5: e1000476.

[9] Weinberger DM, Harboe ZB, Flasche S, Scott JA, Lipsitch M (2011) Prediction of Serotypes Causing Invasive Pneumococcal Disease in Unvaccinated and Vaccinated Populations. Epidemiology 22: 199–207.

[10] Auranen K, Mehtälä J, Tanskanen A, S Kaltoft M (2010) Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. American journal of epidemiology 171: 169–76.

[11] Roberts GO, Rosenthal JS (2001) Optimal scaling for various Metropolis-Hastings algorithms. Statistical Science 16: 351–367.

[12] Auranen K, Mehtala J, Tanskanen A, Kaltoft MS (2010) Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. American journal of epidemiology 171: 169–176.

[13] Miller E, Andrews NJ, Waight PA, Slack MP, George RC (2011) Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. The Lancet Infectious Diseases 11: 760–768.

[14] Choi YH, Jit M, Flasche S, Gay N, Miller E (2012) Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. PLoS ONE 7: e39927.

[15] (2010). Oxford online Dictionary. URL `http://oxforddictionaries.com/`.

[16] Mutsuo Saito, Makoto Matsumoto (2006) SIMD-Oriented Fast Mersenne Twister: a 128-bit Pseudorandom Number Generator. In: Monte Carlo and Quasi - Monte Carlo Methods 2006, Springer. pp. 607–622. doi:10.1007/978-3-540-74496-2_36.

[17] Fine PE (1977) A commentary on the mechanical analogue to the Reed-Frost epidemic model. American journal of epidemiology 106: 87–100.

6 An individual-based model to investigate the role of immunity in the coexistence of the pneumococcal serotypes

Outline

The ecology of *S. pneumoniae* with its distinct features of competition and coexistence is poorly understood. Following the introduction of the 7-valent pneumococcal conjugate vaccine this has become a problem for public health through the observed serotype replacement of NVTs which was initiated by a reduction in competition to these non vaccine serotypes by the VTs (see chapter 3 and [1]). Modelling approaches have been mostly confined to study the inter-dependencies of two (VT/NVT) serotype groups employing models similar to the diamond type model discussed in section 4.3.4. Consequently these include the essential aspect of the transmission dynamics of the pneumococcus -competition- as a rather arbitrary factor which is not motivated by any biological knowledge and introduces caveats to the model predictions (see section 4.3.1).

In section 2.4.3 the immunology of the pneumococcus was studied. Evidence for both serotype specific and serotype non-specific immunity induced by carriage was found. This section studies the potential role of temporarily reduced homologous and heterologous acquisition rates following acquisition of a pneumococcal serotype. To incorporate the complexity of the pneumococcus and to explore the coexistence conditions of multiple serotypes an individual-based model is implemented consisting of up to 20 serotypes. The relationship of this model to

previously assessed deterministic models is investigated and the parameter space for stable coexistence is explored. Finally, the potential implications regarding vaccination in this complex ecology are studied. This research has been submitted for publication.

6.1 Introduction

The emergence of the previously less prevalent serotypes in the absence of the previously dominating vaccine serotypes is likely due to competition between the serotypes [1]. Vaccination against some serotypes is thought to vacate an ecological niche which is filled by serotypes which previously have been outcompeted. The exact source of the competition is not well understood but is believed to arise from protection against additional carriage of other serotypes induced by acquisition of infection [2].

Understanding the transmission dynamics of the pneumococci including the impact and the mechanism of competition between serotypes is essential for predicting the potential impact of future vaccine formulations and therefore is of major importance for public health. However there have been few approaches to describe competition in pneumococcal transmission models and in particular none that did not introduce competition between serotypes as an artificial factor reducing susceptibility against one serotype while carrying another type [3–5]. These approaches to modelling competition were found to artificially increase coexistence [6] and models proposed to incorporate competition in a structurally more meaningful way suffer from a very small parameter space (essentially both serotype groups need to have the same fitness, i.e. $R_{01} = R_{02}$) where coexistence of as few as two serotypes is possible (see chapter 4 and [7]).

The existence of competition raises a more general question about the coexistence of the various pneumococci: In the light of basic evolution dynamics [8] and vast differences in carriage duration of the serotypes (and therefore survival advantages) [9] what is the mechanism ensuring persistence of the variety of pneumococcal serotypes? Apart from possible outbalancing survival advantages like higher transmissibility of the serotypes which are carried for a shorter duration, for which there is no evidence, a decreased likelihood of further colonisation following acquisition of a serotype (serotype-specific immunity) could support coexistence. Data on naturally acquired immunity against pneumococcal carriage

which could possibly support this hypothesis is sparse and the use of biological markers as correlates of protection is not fully understood yet (see section 2.4.3). However, evidence from longitudinal carriage studies was found for both serotype specific and serotype non-specific immunity following acquisition (see section 2.4.3 and [10, 11]).

A structurally simple individual-based model framework to investigate the potential role of both serotype specific immunity and serotype non-specific immunity in the transmission dynamics of *S. pneumoniae* is presented here. The model specifically analyses the likely contribution of immunity to serotype competition and coexistence.

6.2 Methods

6.2.1 Model description

An individual-based model is set up describing the dynamics of up to twenty arbitrary serotypes. Twenty types were chosen as a trade-off between necessary model complexity, computational speed and memory restrictions. At each time step each individual has the attributes: age, immunity duration remaining for each serotype and carriage duration remaining for each serotype. A remaining duration of carriage bigger than zero corresponds to being infected and infectious with the respective serotype. A remaining duration of immunity bigger than zero corresponds to being immune against acquisition of the respective serotype while immunity duration of zero translates to being susceptible to infection with that serotype.

The population demographics are determined by a simple birth and death process: at each time-step a fixed number of susceptible newborns enter the population, population survival rates are applied and the remaining individuals age accordingly. A calendar year in the model was assumed to consist of 12 months and each month of 4 weeks. The population is initialised to its calculated demographic steady state and 1% prevalence of each serotype is allocated amongst the population independently of age. This was found to be a sufficiently large proportion to ensure persistence of sufficiently fit serotypes after initialisation.

Population based mixing patterns were employed. Mixing rates between different

age groups were obtained from a contact survey conducted in the UK [12] and the risk of transmission per contact were used to calculate the age- and serotype-specific force of infection (time dependencies were omitted for clarity):

$$\lambda_{a,s,\Delta} = \sum_{\hat{a}} I_{\hat{a},s} \frac{P^{UK}_{\hat{a}}}{P^{M}_{\hat{a}}} q_{\hat{a},\Delta} \beta_{a,\hat{a}}$$

where $\hat{a}$ is the contact's age group, $I_{\hat{a},s}$ is the number of individuals in age group $\hat{a}$ infected with serotype s, $\beta_{a,\hat{a}}$ the contact rate of individuals in age groups a and $\hat{a}$ and $q_{\hat{a},\Delta} = \hat{q}_{\hat{a}}\sigma_{\Delta}$ the age dependent risk of transmission per contact for scenario S ($\hat{q}_{\hat{a}}$ is the age dependent risk of transmission per contact and $\sigma_{\Delta} \in 1, 2, 4$ representing the low, mid and high FOI scenario). $\frac{P^{UK}_{\hat{a}}}{P^{M}_{\hat{a}}}$ (UK and model population in age group $\hat{a}$ respectively) weights the model population by the real population and ensures transmission according to the age distribution in the UK (compare to section 5.3.2).

At a certain time-step an individual of age a has the probability to be infected with strain s of: $\lambda^{*}_{a,s} = \lambda_{a,s} 1_{R_s}$, where 1_{R_s} is 0 if this individual is currently immune to acquisition of serotype s and 1 otherwise. For an individual of age a let the interval $[0, 1]$ be split up into

$$[0,1] = [0, \lambda^{*}_{a,1}[\ \cup \ [\lambda^{*}_{a,1}, \sum_{i=1}^{2} \lambda^{*}_{a,i}[\ \cup \ [\sum_{i=1}^{2} \lambda^{*}_{a,i}, \sum_{i=1}^{3} \lambda^{*}_{a,i}[\ \cup \ \ldots \ \cup \ [\sum_{i=1}^{20} \lambda^{*}_{a,i}, 1].$$

Whether an individual becomes infected or not is determined by the draw of a uniformly distributed random number $x \in U(0, 1)$ and the interval it is included in. That is $x \in [\sum_{i=1}^{j-1} \lambda^{*}_{a,i}, \sum_{i=1}^{j} \lambda^{*}_{a,i}[$ means infection with serotype j and $x \in [\sum_{i=1}^{20} \lambda^{*}_{a,i}, 1]$ corresponds to no acquisition of carriage. Note that for this $\sum_{i=1}^{20} \lambda^{*}_{a,i} \leq 1$ is required which is checked throughout all simulations. This formulation only allows for the acquisition of a maximum of one serotype at each time time-step, which implicitly assumes instant protection following acquisition of carriage.

Following acquisition an individual is assigned a duration of carriage, length of immunity duration against repeated acquisition with the same serotype (specific immunity) and length of immunity against acquisition of any other serotype (non-specific immunity). Specific and non-specific immunity are assumed to offer full protection against acquisition of infection of the same and of all types respectively.

Therefore if the sampled duration of non-specific immunity exceeded the specific one (the case that specific immunity offers no additional protection) protection against acquisition of colonisation by this specific serotype was assumed to last for the duration of the non-specific immunity. If the duration of carriage exceeds serotype non specific (specific) immunity additional heterologous (homologous) acquisition of carriage is possible. Additional homologous acquisition is modelled through an accordingly increasing duration of carriage. Similarly, if an individual has remaining specific immunity against one type, and then acquires additional heterologous carriage, the duration of immunity against the resident serotype is extended by non-specific immunity.

For each serotype each individual is in any of the following states: not colonised & susceptible, not colonised & immune, colonised & susceptible, colonised & immune (so more than $4^{20} \approx 10^{12}$ possible states per individual overall). The limiting factor, which prevents many of these states, is that existing immunity prevents further acquisition. The immunity considered in this model is only short-term immunity offering protection directly after acquisition of colonisation. However, maturation of the immune system and increasing serotype non-specific immunity from repeated exposure is accounted for through decreasing risk of transmission per contact and carriage duration by age.

The model simulating the pneumococcal ecology before the introduction of any vaccination was run to simulate 30 years. Although after 5 years a stable equilibrium was reached for most simulations, for some parameterisations at the extremes of the considered parameter space this took significantly longer, and a burn in period of 20 years was allowed for. Hence, the results presented in the following are based on 10 years of simulation at steady state. For comparability amongst different serotypes carriage prevalence was standardised by the respective mean (standardised prevalence) when studying serotype volatility.

To simulate a conjugate vaccine offering protection against a limited amount of serotypes and a seemingly ideal vaccine offering the same level of protection against all serotypes, two different vaccine scenarios were considered: a) a vaccine providing immunity against the two most prevalent types (bi-valent vaccine) and b) a vaccine providing immunity against all serotypes (universal vaccine). In both cases the underlying parameters were the same. The model was allowed a burn-in period of 20 years. After 30 years either of the vaccines were introduced: all children at the age of 2 months received one dose which offered a 65% chance

for a ten year protection against the acquisition of the serotypes included in the respective vaccine formulation [13]. The period between year 20 and 30 is used as a pre-vaccination baseline and the period between year 50 and 60 is used as the post vaccination period where vaccine effects have stabilised.

Further the impact of extending an existing bi-valent vaccination program to a universal vaccine was investigated. Generally, no importation of infection was considered as no migration is assumed in the model, i.e. when there were no infections with a serotype at any point then it is extinct for the remaining course of the simulation. However, for this particular impact investigation each of the twenty serotypes were allowed to re-enter the population each month. Therefore one additional infectious individual per type was introduced which allowed for re-emergence of previously extinct serotypes.

6.2.2 Parameter assumptions

An overview of model parameters is provided in table 6.1. The model was run in two, four or eight time steps per week for the low, mid and high FOI scenario respectively to ensure $\sum_{i=1}^{20} \lambda_{a,i}^{*} \leq 1$ at all times. A total of 64 newborns enter the population each week to achieve an average population size of 243,792 individuals. Age depended mortality rates up to 80 years of age were available from the Office for National Statistics in 2007 and transformed to survival rates s_a. For the age cohort of individuals aged between 80 and 100 years old an exponentially decaying survival rate was assumed with an assumed maximum possible age of 100 years ($s_{100} = 0$). The corresponding survival rates per time-step of an individual at age a (years) follow as $s_a^{\frac{1}{hT}}$, with h being the number of time-steps per week and T the number of weeks per year.

The assumption was made that while a person is infected with the pneumococcus he is infectious and that infectiousness does not vary over the course of the infection. Carriage duration was assumed to be the number of weeks where the infection could not be cleared by the immune system and therefore to follow a negative binomial distribution. In the absence of further knowledge no overdispersion was assumed and so the variance was assumed to be equal to the mean. The mean duration of carriage in children younger than 2 years of age was set gradually decreasing from 8 weeks for serotype 1 to 3.25 weeks for serotype 20 representing differences in fitness. 2-4 year old children were assumed to be infected

Table 6.1: An overview of the essential parameter values and assumption in the simulation model

Parameter	**Description**	**Value**	**Source**
δ_N	mean duration of non specific immunity	$\in 1, 2, \ldots, 40$ weeks	exploration of parameter space
δ_S	mean duration of specific immunity	$\in \{1, 2, \ldots, 52\}$ weeks	exploration of parameter space
σ_S	transmission modifier (low, mid, high FOI)	$\in \{1, 2, 4\}$	exploration of parameter space
a	transmission age groups in years	$[0, 2[, [2, 5[, [5, 10[, [10, 20[, [20, 40[, [40, \infty[$	
$\hat{q}_a$	age dependent risk of transmission per contact	0.033, 0.030, 0.026, 0.018, 0.011, 0.007	to mirror age decline in [14]
$\beta_{1,\hat{a}}$	contact rate of 0-1 year olds with other age groups	$(41.7, 32.6, 11.0, 5.6, 12.8, 4.2)10^{-8}$/day	Mossong et al. [12]
$\beta_{2,\hat{a}}$	contact rate of 2-4 year olds with other age groups	$(32.6, 105, 26.7, 9.8, 18.3, 7.5)10^{-8}$/day	Mossong et al. [12]
$\beta_{3,\hat{a}}$	contact rate of 5-9 year olds with other age groups	$(11.0, 26.7, 110, 13.9, 13.9, 6.1)10^{-8}$/day	Mossong et al. [12]
$\beta_{4,\hat{a}}$	contact rate of 10-19 year olds with other age groups	$(5.6, 9.8, 13.9, 53.4, 8.8, 6.2)10^{-8}$/day	Mossong et al. [12]
$\beta_{5,\hat{a}}$	contact rate of 20-39 year olds with other age groups	$(12.8, 18.3, 13.9, 8.82, 13.7, 6.9)10^{-8}$/day	Mossong et al. [12]
$\beta_{6,\hat{a}}$	contact rate of 40+ year olds with other age groups	$(4.2, 7.5, 6.1, 6.2, 6.9, 6.8)10^{-8}$/day	Mossong et al. [12]
ω_1	mean duration of carriage in 0-1 year olds	$3.25, 3.5, \ldots, 8$ weeks	Melegaro et al. [15]
ω_2	mean duration of carriage in 2-4 year olds	$1.625, 1.750, \ldots, 4$ weeks	Melegaro et al. [15]
ω_3	mean duration of carriage in 5+ year olds	$0.8125, 0.8750, \ldots, 2$ weeks	Melegaro et al. [15]
s_a	yearly survival rates in E&W in 2007		Office for National Statistics

only for half and individuals >5 years for quarter of that duration [5, 9].

Carriage duration is the only determinant for differences in fitness of the pneumococcal serotypes in the model. Carriage duration being a major component determining fitness, and therefore prevalence of the pneumococci, is supported by Weinberger et al. [16]. These authors found that the capsule size of the pneumococcal serotypes is correlated with both avoidance of neutrophil-mediated killing and prevalence. However other factors, such as possible differences in transmissibility, could also influence a serotype's fitness but are disregarded in the model.

Durations of non-specific and specific short-term immunity were assumed to follow Negative Binomial distributions for the same reason as with the duration of carriage and their potential mean values were explored on a grid $[0, 40]$ weeks$\times[0, 52]$ weeks. Non-specific immunity is assumed to allow instant protection against all pneumococci. Allowing for a slight delay until activation would result in an increased time span for possible multiple acquisition and reduced competition. A delay in the protection offered by specific immunity was not included, because any delay in specific protection was assumed to be compensated for by the non-specific immune response.

Data on two-way conversational contact patterns from a contact survey including the UK were used to calculate normalised age-specific contact rates for the age groups 0-1, 2-4, 5-9, 10-19, 20-39, 40+ [12]. The risk of transmission per contact was chosen to decrease by age in order to mirror the decrease in prevalence with age [9] and to cause the reproduction number R_0 to be in the range of 1.1 to 2.7 for the different serotypes (low FOI scenario), 2.2 to 5.4 (mid FOI scenario) or 4.4 to 10.8 (high FOI scenario). Carriers of multiple pneumococci are assumed to be as transmissive with each type as a single carrier is. This may overestimate the FOI but the effect of multiple colonisation on transmission is yet to be studied and the effect should be small if less than 10% of infected individuals are colonised with another type, as is found in most studies to date. However this might be altered with the emergence of new detection techniques (see section 2.5.2).

6.2.3 Assessment of cyclic behaviour

The robustness of the parameter space against cyclical behaviour of the pneumococcus was determined. Suppose that the time course of prevalence of a serotype can be described by a stationary autoregression model of order 2, $X_t =$

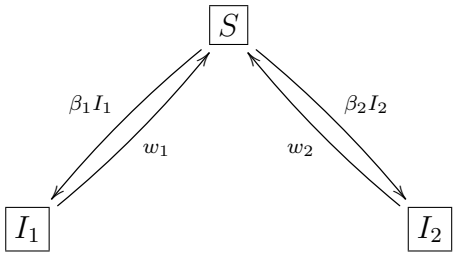
$\phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$. If the solutions of its characteristic polynomial are complex numbers the autocorrelation function of X_t would be a combination of sine and cosine waves and the prevalence would show cyclic behaviour [17]. This is the case if $\phi_1^2 + 4\phi_2 < 0$. Therefore the prevalence of a serotype was defined to be cyclic/unstable if after the model reached the steady state the above condition is met and the coefficient of variation exceeded 0.15. This value is arbitrary, but probably reflects a reasonable threshold from which similar patterns might have been detected in longitudinal carriage studies.

6.2.4 Comparison to deterministic modelling approaches

To visualise various parameterisations of the simulation model a comparison to some simple deterministic approaches is carried out. This shows how carriage duration and specific and non-specific immunity duration influence the structure of the respectively parameterised simulation model. Five deterministic 2-strain models were defined which could easily be extended to incorporate a simple birth and death process and an age structure to reflect the approach taken in the IBM. These models range from a simple separable model to a diamond type model:
Model 0:

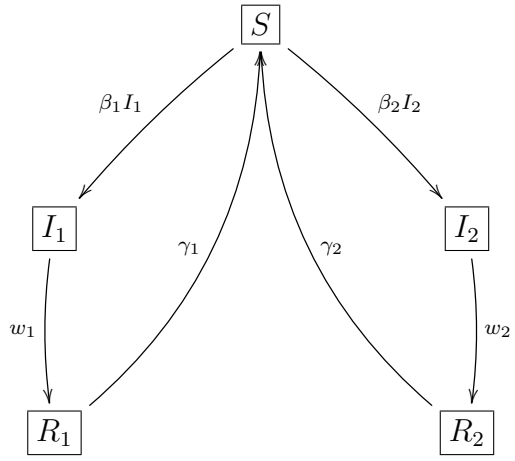
$$
\begin{aligned}
S_1' &= -\beta_1 S_1 I_1 + w_1 I_1 \\
I_1' &= +\beta_1 S_1 I_1 - w_1 I_1 \\
S_2' &= -\beta_2 S_2 I_2 + w_2 I_2 \\
I_2' &= +\beta_2 S_2 I_2 - w_2 I_2
\end{aligned}
$$

Model 1:



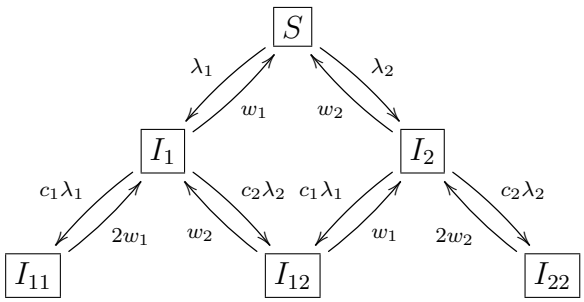
$$S' = -\beta_1 S I_1 - \beta_2 S I_2 + w_1 I_1 + w_2 I_2$$
$$I_1' = +\beta_1 S I_1 - w_1 I_1$$
$$I_2' = +\beta_2 S I_2 - w_2 I_2$$

Model 2:



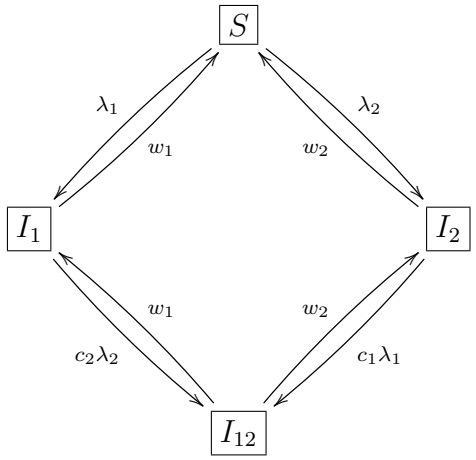
$$S' = -\beta_1 S I_1 - \beta_2 S I_2 + \gamma_1 R_1 + \gamma_2 R_2$$
$$I_1' = +\beta_1 S I_1 - w_1 I_1$$
$$I_2' = +\beta_2 S I_2 - w_2 I_2$$
$$R_1' = +w_1 I_1 - \gamma_1 R_1$$
$$R_2' = +w_2 I_2 - \gamma_2 R_2$$

Model 3:



$$
\begin{aligned}
S' &= -(\lambda_1 + \lambda_2)S + w_1 I_1 + w_2 I_2 \\
I_1' &= +\lambda_1 S - w_1 I_1 - c_2\lambda_2 I_1 + w_2 I_{12} - c_1\lambda_1 I_1 + 2w_1 I_{11} \\
I_2' &= +\lambda_2 S - w_2 I_2 - c_1\lambda_1 I_2 + w_1 I_{12} - c_2\lambda_2 I_2 + 2w_2 I_{22} \\
I_{11}' &= +c_1\lambda_1 I_1 - w_1 I_{11} \\
I_{12}' &= +c_2\lambda_2 I_1 + c_1\lambda_1 I_2 - (w_1 + w_2) I_{12} \\
I_{22}' &= +c_2\lambda_2 I_2 - w_2 I_{22} \\
&\text{with } \lambda_1 = \beta_1(I_1 + I_{12} + 2I_{11}) \text{ and } \lambda_2 = \beta_2(I_2 + I_{12} + 2I_{22})
\end{aligned}
$$

Model 4:



$$
\begin{aligned}
S' &= -(\lambda_1 + \lambda_2)S + w_1 I_1 + w_2 I_2 \\
I_1' &= +\lambda_1 S - w_1 I_1 - c_2\lambda_2 I_1 + w_2 I_{12} \\
I_2' &= +\lambda_2 S - w_2 I_2 - c_1\lambda_1 I_2 + w_1 I_{12} \\
I_{12}' &= +c_2\lambda_2 I_1 + c_1\lambda_1 I_2 - (w_1 + w_2) I_{12} \\
&\text{with } \lambda_1 = \beta_1(I_1 + I_{12}) \text{ and } \lambda_2 = \beta_2(I_2 + I_{12})
\end{aligned}
$$

Similar to the naming conventions employed earlier (see chapters 0 and 4) S, S_1, S_2 represent the part of the population who are not carrying both strains or either one strain, I_1, I_2 the part who are carrying either one of the strains and I_{11}, I_{22}, I_{12} the part who are carrying two copies of either strain or both strains. $\beta_{1,2}$ are the type-specific contact rates, $w_{1,2}$ the clearance rates, $\gamma_{1,2}$ the rates of immunity loss after clearance of colonisation, $\lambda_{1,2}$ the forces of infection and $c_{1,2}$ the competition parameters. If the individual-based model presented here is set to only consist of two strains (i.e. set to initial prevalence for the remaining 18 strains to 0) the transmission dynamics incorporated for specific parameter choices are similar to

the deterministic models presented.

0. If the duration of serotype specific immunity is fixed to be equal to the duration of carriage and if the duration of non-specific immunity is set to zero the individual-based model corresponds to model 0. Individuals are immune against re-acquisition of the colonising serotype exactly for their duration of carriage and the absence of non-specific immunity induces completely independent transmission of both serotypes.

1. When the duration of carriage is equal to the duration of non-specific immunity and specific immunity offers no additional protection (that is if it is shorter or equal to the duration of non-specific immunity) both serotypes have to compete for the same pool of susceptibles and can only be colonised by one serotype at a time. This model corresponds to model 1 (the model discussed in section 4.3.2) and it can be easily shown that coexistence of both serotypes in this model is only possible for the case that both serotypes share the same reproduction number.

2. The IBM corresponds to model 2 if the duration of non-specific immunity is longer than the duration of carriage in the absence of additional protection from serotype specific immunity. This yields a new possible state where an individual can be no longer infectious but still immune against acquisition of any pneumococci.

3. Still in the absence of additional protection from specific immunity and in the case that the duration of non-specific immunity is shorter than the duration of carriage an individual can be co-infected by an additional strain. This includes co-infection with the same strain (co-infection with the same strain in the IBM is realised through adding up the duration of carriage and immunity). Hereby the competition factors (c_1, c_2) in the deterministic model (model 3) work similarly to the IBM where competition arises through a reduced period where acquisition is possible while carrying. The shorter that period the stronger the competition.

4. When the duration of non-specific immunity is shorter than the duration of carriage which itself is shorter than the duration of serotype specific immunity then the transmission dynamics in the IBM are similar to model 4 (also see section 4.3.4). The long duration of strain specific immunity prohibits co-colonisation with the same serotype while a short duration of

non-specific immunity permits co-colonisation with other serotypes.

In the individual-based model the durations of non-specific immunity, specific immunity and carriage were sampled independently from a Negative Binomial distribution for each newly acquired infection. Therefore the infection process of the IBM is a mixture of all these deterministic models but the choice of the mean values for the distributions determines the proximity to the various deterministic models.

6.2.5 Justification for the use of carriage duration as the main determinant for differences in fitness of the pneumococci

One of the model assumptions is that differences in the duration of carriage (or rather the susceptibility to clearance) is the main determinant for carriage prevalence and therefore acts like a representative of a serotype's fitness. This is supported by Weinberger et al. [16] who found that heavily encapsulated pneumococci are more resistant to neutrophil-mediated killing, which increases their duration of carriage. Further, these authors establish a relation between encapsulation and carriage prevalence and conclude that while highly prevalent types tend to be heavily encapsulated less prevalent types are not necessarily less encapsulated. That is types with a thin capsule are generally not prevalent amongst pneumococcal carriage, but for types with a thick capsule one can't really predict their contribution to carriage.

There are two ways how the length of carriage duration can influence carriage prevalence: (i) longer duration of carriage makes types more likely to be picked up in cross sectional studies and (ii) the extended duration of carriage allows for a longer period where transmission could take place (increase in acquisition events). Since the first one is obvious only the second one needs further investigation, and hence data presented in Sleeman et al. [18] are analysed. These consist of nasopharyngeal carriage episodes amongst children under the age of 2 years and reports the duration of carriage and carriage incidence (acquisition of infection). By adjusting carriage prevalence for the duration of carriage the correlation of duration of carriage and acquisition events can be estimated. If carriage duration was the only determinant of carriage prevalence (for example if transmission of the pneumococcus could only occur shortly after acquisition - as is the case for

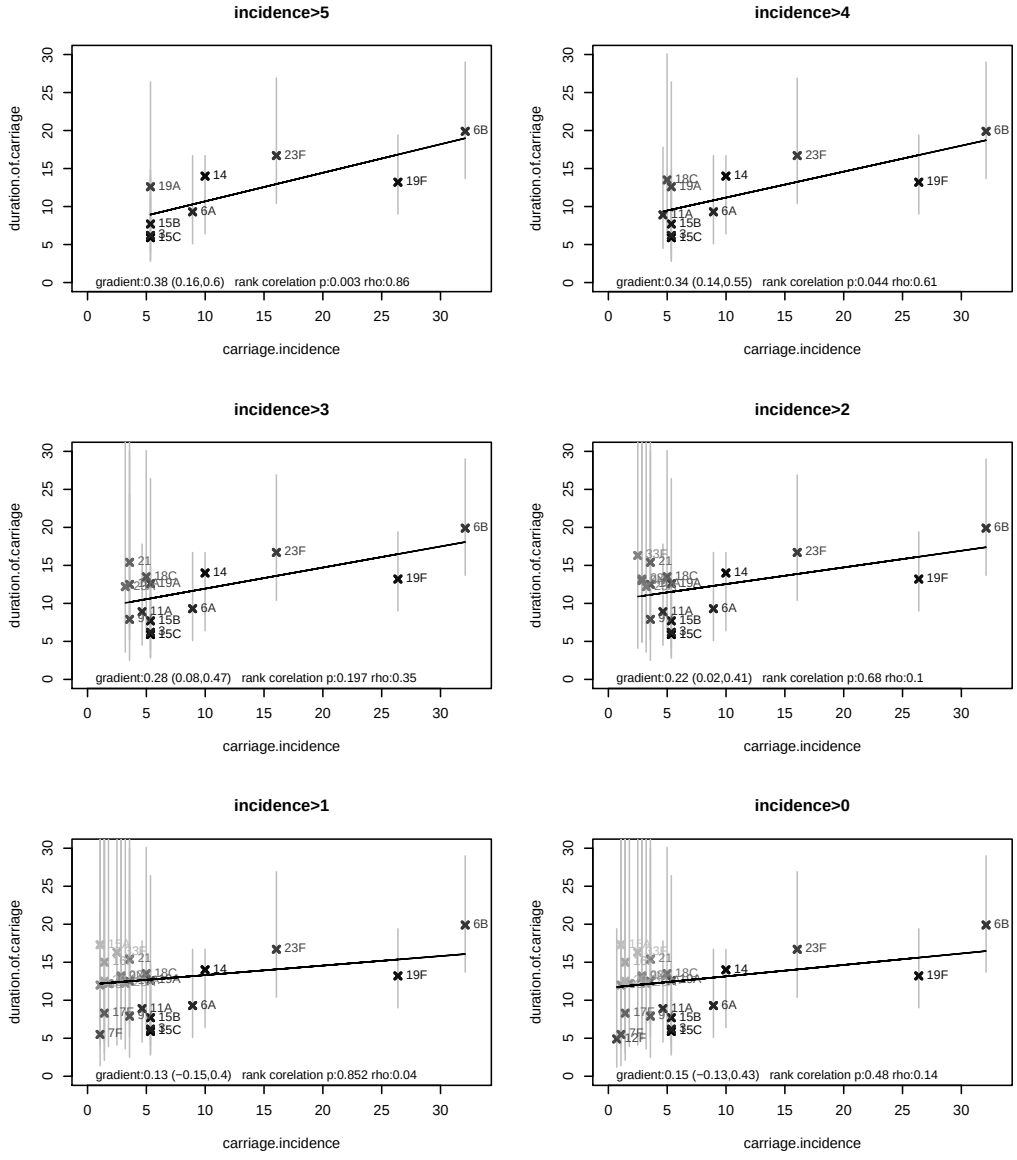


Figure 6.1: Serotype specific carriage duration (weeks) over carriage incidence (number of pneumococcal acquisitions per 100,000 child years). Data with confidence bounds (CI) on carriage duration as from Sleeman et al. [18] are presented together with a regression line. Serotypes are restricted by their incidence being bigger than 5,4,3,2,1,0 respectively. The ranking of the length of the confidence intervals is represented by the colour saturation, where dark grey corresponds to a narrow CI and light grey wide CI.

progression to invasive disease) then no correlation between carriage duration and carriage incidence should exist.

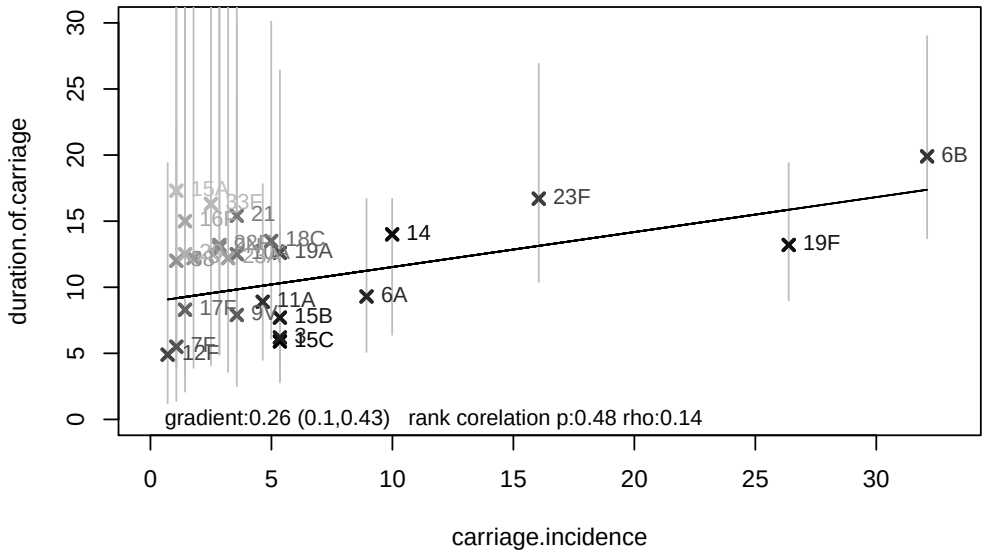


Figure 6.2: Serotype specific carriage duration (weeks) over carriage incidence (pneumococcal acquisition). Data with confidence bounds (CI) on carriage duration as from Sleeman et al. [18] are presented together with a the model fit of a weighted linear regression. The ranking of the length of the confidence intervals is represented by the colour, where dark grey corresponds to a narrow CI and light grey wide CI.

A simple linear regression and Spearman's rank correlation test were computed to investigate whether there is correlation between carriage duration and acquisition of infection. Employing the point estimates for both carriage duration and incidence for all serotypes no significant correlation can be detected (see figure 6.1). However, when restricting the analysis to those types who are sufficiently present (incidence higher then 4 acquisitions per 100,000 child years and therefore less variance in the point estimate of incidence) a significant positive correlation becomes evident. This is also apparent when carrying out a weighted linear regression instead, where the inverse standard errors of the estimates on the duration of carriage are employed as weights (see figure 6.2).

Overall, evidence supporting carriage duration as the main determinant for fitness and hence dominance in carriage prevalence is weak but may possibly be a reasonable assumption in the absence of knowledge of other contributors. On the basis of data on carriage incidence this section further suggests a possible implicit

role for the duration of carriage in transmission of the pneumococcus.

6.3 Results

6.3.1 Competition and multiple carriage

The modelled serotypes compete with each other through infection-induced immunity against acquisition of all serotypes (non-specific immunity), which leads to a reduced pool of susceptibles to be infected by other serotypes; so with increasing duration of non-specific immunity fewer serotypes can coexist due to increasing competition (figures 6.3 and 6.4). The difference in duration of carriage and the duration of non-specific immunity (in combination with the force of infection) thereby determines the rate of multiple carriage in the population; a relatively short duration of non-specific immunity does allow for additional infection(s) whereas the case of non-specific immunity being longer lasting than carriage denies other serotypes the ability to infect the host at the same time, see figures 6.3 and 6.4.

6.3.2 Coexistence

Exploring the parameter space of specific and non-specific immunity yields a strong dependence of coexistence on specific immunity (see figures 6.3 and 6.4). If serotypes limit their own spread in the same way as they limit the spread of competing serotypes serotype coexistence is hardly possible; in fact if non-specific immunity duration is smaller than or equal to specific immunity duration one serotype only will dominate the others and lead to their extinction. This can be seen in the upper panel of figure 6.3 where each line representing the number of coexisting serotypes for a specific duration of non specific immunity only raises above one once the duration of specific immunity exceeds the duration of non specific immunity. In the model the immunity durations are drawn from a Negative Binomial distribution, therefore only the relationship "mean non-specific immunity duration $\leq$ mean specific immunity duration" holds which allows for a few serotypes to coexist in the scenarios of medium and high force of infection (Figure 6.5). However, with increasing duration of serotype specific immunity, the numbers of serotypes coexisting stably increases and especially for the mid-

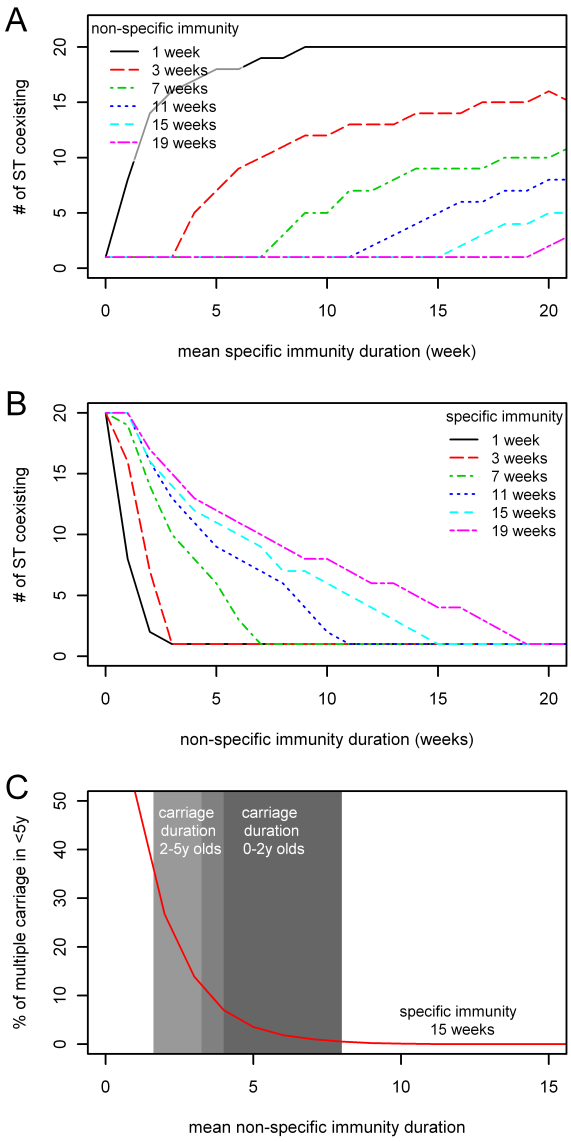


Figure 6.3: The impact of specific immunity on coexistence, for non-specific immunity and carriage duration on coexistence and multiple carriage for the low FOI scenario. This presents the results after 30 years of simulation for one simulation at each combination of specific and non specific immunity. Only durations shorter than 20 weeks are presented since for longer durations coexistence is influenced by cyclic behaviour (compare section 6.3.5). A) Positive correlation of the number of coexisting serotypes and the mean duration of specific immunity for scenarios of a fixed mean duration of non-specific immunity. B) Negative correlation of the number of coexisting serotypes and the mean duration of non-specific immunity for scenarios of a fixed mean duration of specific immunity. C) Declining proportion of multiple colonisation amongst the <5 year olds with increasing duration of non-specific immunity (specific immunity fixed to 15 weeks) set in comparison with the average duration of carriage. The two overlapping grey areas indicate the assumed duration of carriage in relation to the mean specific immunity duration.

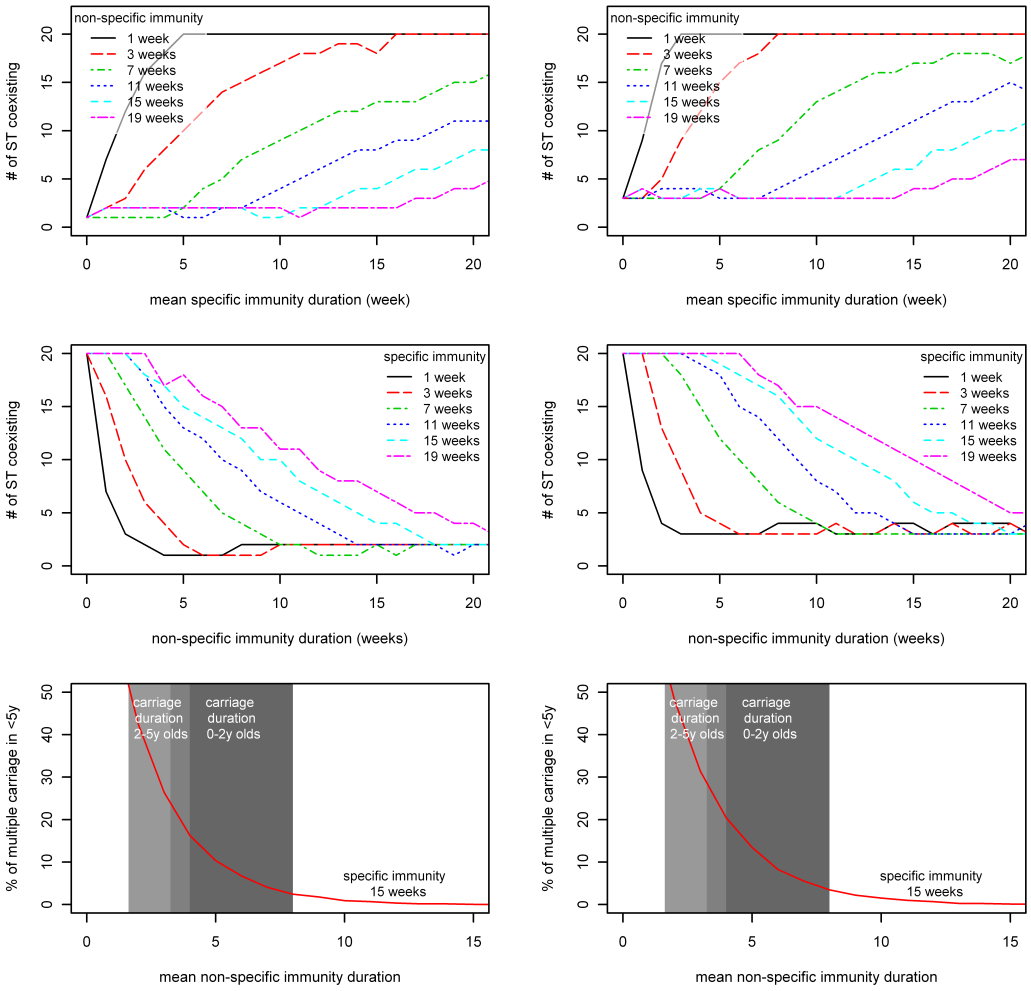


Figure 6.4: The impact of specific immunity on coexistence, for non-specific immunity and carriage duration on coexistence and multiple carriage. Corresponding to figure 6.3 this shows the results for the mid (left) and high (right) FOI scenario.

and high- FOI scenarios there is a broad parameter space where more than 90% of all serotypes coexist (Figure 6.5).

6.3.3 Variance of low prevalence serotypes

The variance of the standardised carriage prevalence was found to increase with decreasing prevalence level as expected. This means that the serotypes at the verge of extinction show high variance. The corresponding peaks can stretch over multiple years (Figure 6.6). If low prevalence types are highly invasive, then this would result in apparent epidemics of invasive disease.

6.3.4 Carriage prevalence and the force of infection

Carriage prevalence generally increases for endemic pathogens with increasing FOI (compare section 1.4.2 on inferring R_0 from the proportion susceptible at equilibrium). However, in contrast to single pathogen models, in this multi serotype model the higher the FOI the higher the competition introduced by other serotypes which leads to only moderate changes in prevalence. Figure 6.7 shows that despite a four fold increase in FOI from the low to the high scenario the effects on overall carriage in <5y olds are only moderate (about 10% increase). The main determinant for carriage prevalence is rather the duration of specific and non-specific immunity.

6.3.5 Parameter space for coexistence

The aim of this section is to explore the influence of the human immune defence system on the versatility of the pneumococcus. The resulting model estimates should not be understood as real world estimates but rather in relation to this model, as comparative estimates for comparing simulation runs. However, some structural conclusions on the likely parameter space which yield coexistence of multiple serotypes in the population could be drawn (see Figure 6.5).

As identified from deterministic modelling approaches [7] it was found that in the absence of additional serotype specific protection through serotype specific immunity (specific immunity is of equal or shorter duration than non-specific

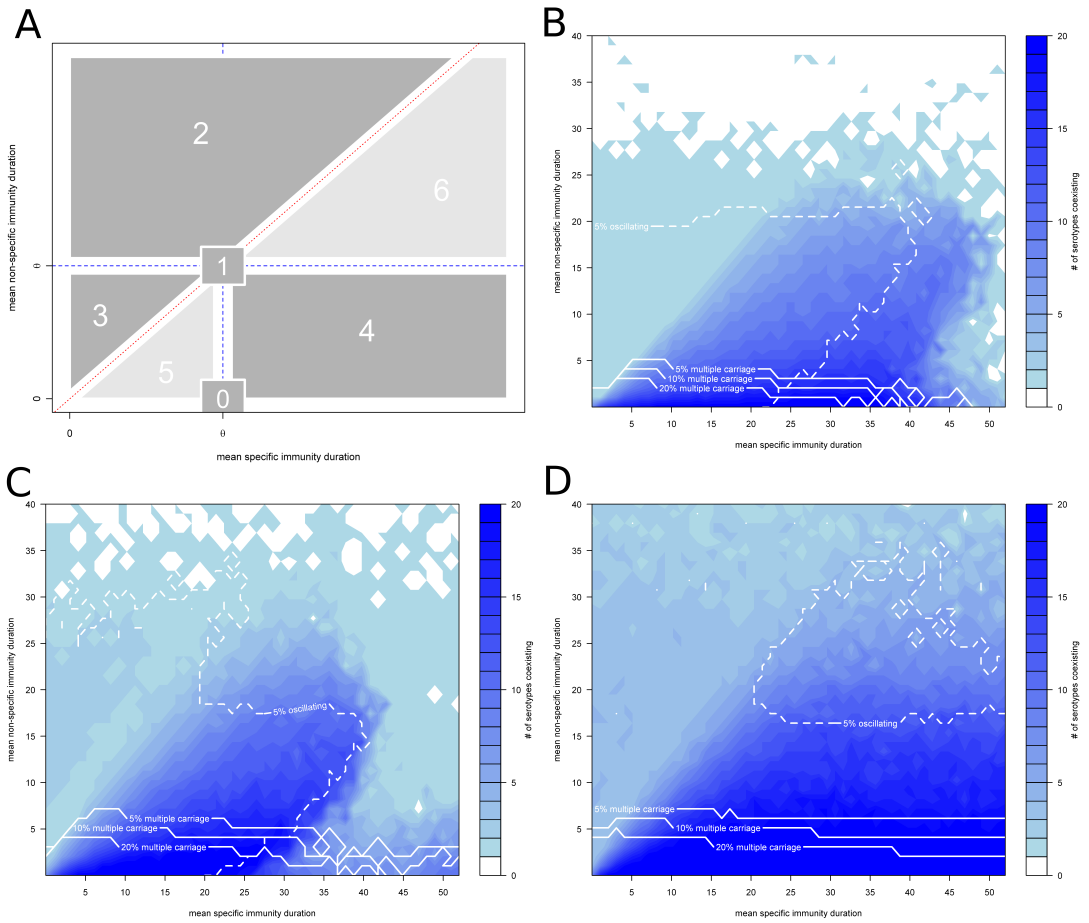


Figure 6.5: Exploration of the parameter space. A) Division of the parameter space of non-specific and specific immunity duration according to proximity to the deterministic models 0 to 4 depending on the duration of carriage $\theta = \frac{1}{w}$ (blue dashed lines). The red line indicates the the same duration of specific and non-specific immunity. The space labelled 5 represents a mixture of model 3 and 4 where individuals benefit from additional serotype specific protection which vanishes before clearance and therefore, even though less likely than heterologous carriage, additional acquisition of homologous carriage is a frequent outcome. The parameter space labelled 6 represents a mixture of model 2 and 4 where individuals benefit from additional serotype specific protection but multiple acquisition is an infrequent event. B-D) The number of stably coexisting serotypes in the population in dependence of the FOI (low B, mid C, high D) and duration of non-specific and specific immunity. The parameter space of high fractions of the population colonised with more than one serotype at a time is indicated by white solid lines. The dashed line marks the space where at least 5% (at least 1 serotype, typically the most prevalent one) of all serotypes shows cyclic behaviour induced by long duration of immunity. In all three scenarios this happens in the parameter space where the duration of specific or non specific immunity is long.

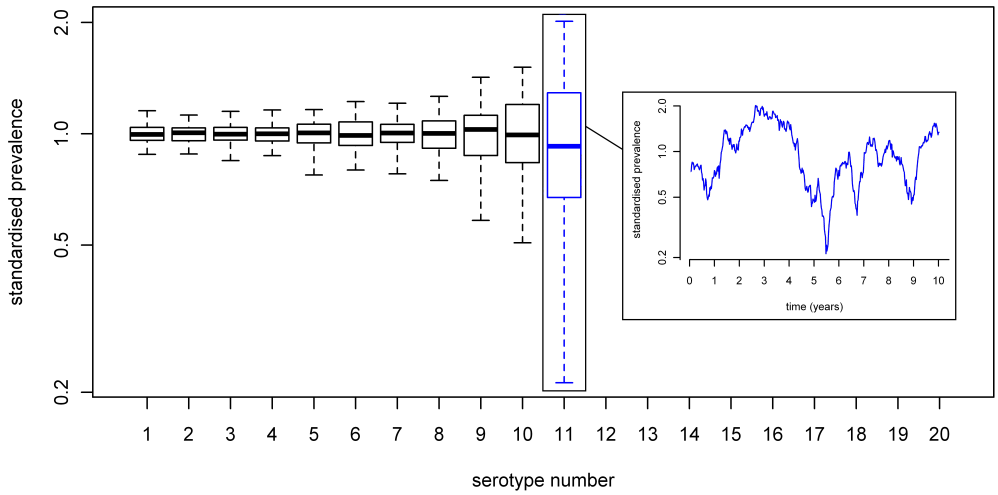


Figure 6.6: A typical example of the variance of the serotypes standardised prevalence after the burn-in period (here, mid FOI scenario; mean specific immunity duration: 18 weeks, mean non-specific immunity duration: 10 weeks, 11 serotypes were found to stably coexist in this scenario). Serotypes at the verge of extinction show relatively high deviation from their mean prevalence and peaks and troughs may stretch over several years.

immunity) no coexistence is possible and the competitive pressure induced by the fittest type leads to extinction of all others. Note again, that for the model simulations immunity durations for both non-specific and specific immunity are drawn from a Negative Binomial distribution, therefore when exploring the parameter space only the mean specific immunity duration is smaller or equal to the mean non-specific immunity duration. This allows for a few serotypes to coexist in this parameter space in the scenarios of medium and high force of infection (see figures 6.4 and 6.5).

As shown earlier in figure 6.3 a negative correlation between the number of coexisting serotypes and non-specific immunity duration generally holds. However, there are limitations to this.

- The difference in duration of non-specific immunity and carriage determines the length of time for which an individual is susceptible for additional acquisition of infection. Hence it determines the amount of multiple carriage in the population (see figures 6.3 and 6.4 and figure 6.5). Therefore relatively short non-specific protection would yield high levels of multiple carriage which have not been observed in data for E&W, even with newly developed

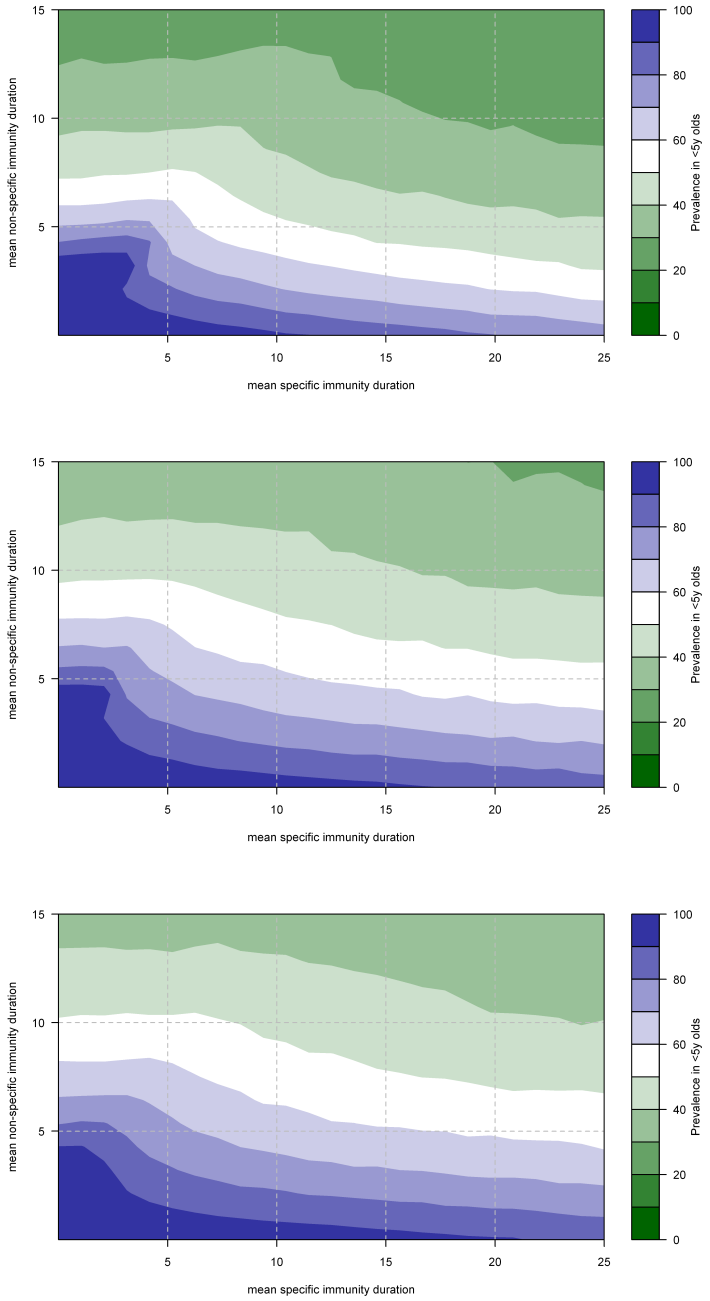


Figure 6.7: Comparison of the effects of immunity duration (in weeks) and force of infection on carriage prevalence in children less than 5 years of age. Low, mid and high FOI scenarios are shown from top to bottom. Note that where duration of specific immunity does not exceed the duration of non specific immunity it doesn't have much of an effect and hence carriage prevalence is similar. When it exceeds non specific immunity the specific immunity limits overall prevalence. While a four-fold change in FOI has only minor impact, the prevalence levels depend hugely on the duration of specific and in particular non-specific immunity.

detection methods rates of multiple carriage in infants in a high prevalence setting were below 50% [19].

- Long duration of non-serotype specific immunity in the extreme will lead to a Susceptible-Infective-Immune - type model behaviour with one single epidemic occurring which eventually leads to a depletion of susceptibles and no infected persons in the population. In less extreme cases frequently occurring cyclic patterns in carriage prevalence can be the result (see figure 6.8) which have not been observed in pneumococcal nasopharyngeal carriage.

Also positive correlation between the number of coexisting serotypes and serotype specific immunity was found. Similar restrictions to the ones in non-specific immunity apply here and impose restrictions to the feasible parameter space:

- The lack of additional type-specific protection means that coexistence of multiple types becomes virtually impossible.

- Long serotype specific protection can cause extinction of serotypes when it induces a long-term lack of susceptibles to become re-infected.

The cyclic behaviour induced by long duration of immunity is most pronounced in the most prevalent types. In figure 6.5 its potential restriction on the parameter space to be considered reasonable for the model is shown.

6.3.6 Model neutrality

As discussed in section 4.3.1 a model is termed a "neutral null-model" if it meets two criteria: 1) "ecological neutrality" - if the strains are indistinguishable the dynamics of the ecological variables (the number of individuals infected with 0,1,2,... serotypes) should only depend on the ecological variables, in particular be independent of specific strains and 2) "population genetic neutrality" - no stable equilibrium should exist i.e. the strain prevalence should be dependent on the initial conditions and it should be possible to initialise indistinguishable strains to different prevalence without having them converge to the same equilibrium (see [6] and section 4.3.1).

The simulation model does not formally meet the ecological neutrality criterion because homologous reinfection is not included as additional carriage of another copy of this serotype. Hence, heterologous reinfection is treated differently to ho-

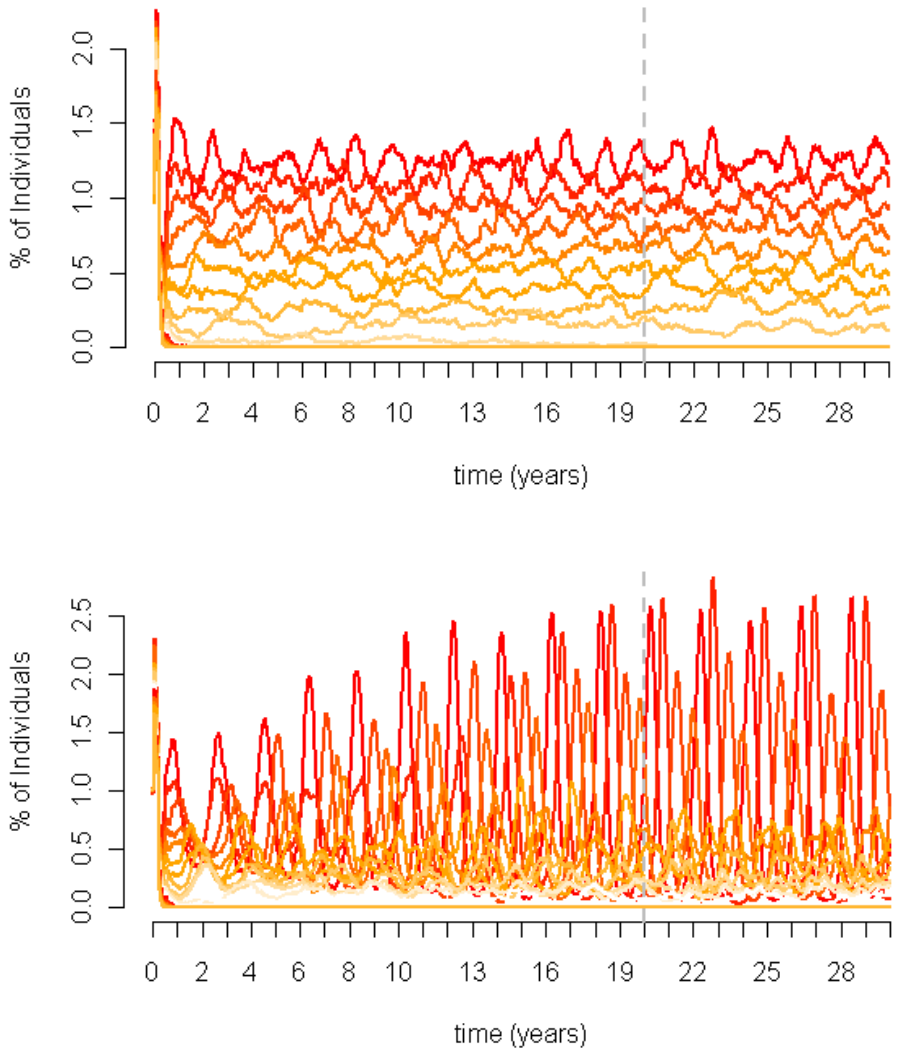


Figure 6.8: Proportion of serotype specific carriage prevalence in the simulation model. The simulation starts with 1% prevalence per type and the burn in period is considered to be 20 years (grey dashed line). Serotypes are coloured according to fitness where dark red corresponds to the longest duration of carriage. The upper panel shows a simulation where specific and non specific immunity duration are 8 and 19 weeks and the lower panel a simulation with 8 and 35 weeks respectively. While the scenario with longer protection from serotype specific carriage shows a greater amount of coexisting serotypes serotypes follow strong cyclic patterns which considerably exceed the stochastic fluctuations in the scenario with shorter duration of type specific immunity.

mologous reinfection and therefore hinders the representation of the dynamics of the ecological variables in dependence of only the ecological variables. However, in order to simulate similar dynamics it was assumed that homologous acquisition of carriage results in an extension of the duration of carriage by another full duration of carriage for this serotype in the respective age group. So, while an individual is not twice as infectious with a serotype following homologous acquisition, and hence the ecological criterion is not met, the individual is infectious for twice the time, on average.

Serotype specific immunity induces a period of protection against a single type whilst the individual is already susceptible against acquisition of infection by other types. Hence the probability of acquiring carriage for this individual is specifically dependent on the identity of this particular type which violates the criterion for ecological neutrality. However, this is conditional on the existence of serotype specific immunity and no such violations arise from serotype non specific immunity.

Population genetic neutrality was assessed for the individual-based model parameterised to simulate the deterministic models 3 and 4 discussed earlier. The simulations are initialised with five serotypes at 1%, 2% 3%,4% and 10% prevalence, respectively. If the prevalence of all serotypes after 30 simulated years converged to the same equilibrium despite their different prevalence at initialisation the population genetic neutrality criterion was considered to be violated. The low FOI scenario was employed using a mean duration of carriage in children younger than 2 years of 8 weeks for all types to make them indistinguishable in fitness. To reflect model 3 the duration of serotype non-specific immunity was set to be 2 weeks less than duration of carriage and no extra protection from serotype specific immunity was assumed. To reflect model 4 the duration of serotype non-specific immunity was set to be 2 weeks less than duration of carriage and serotype specific immunity was set to last two weeks longer than carriage. The population genetic neutrality was found to hold for the IBM parameterised to describe model 3 and to be violated for the representation of model 4 (see figure 6.9). These results mirror the analytic results for the respective deterministic models.

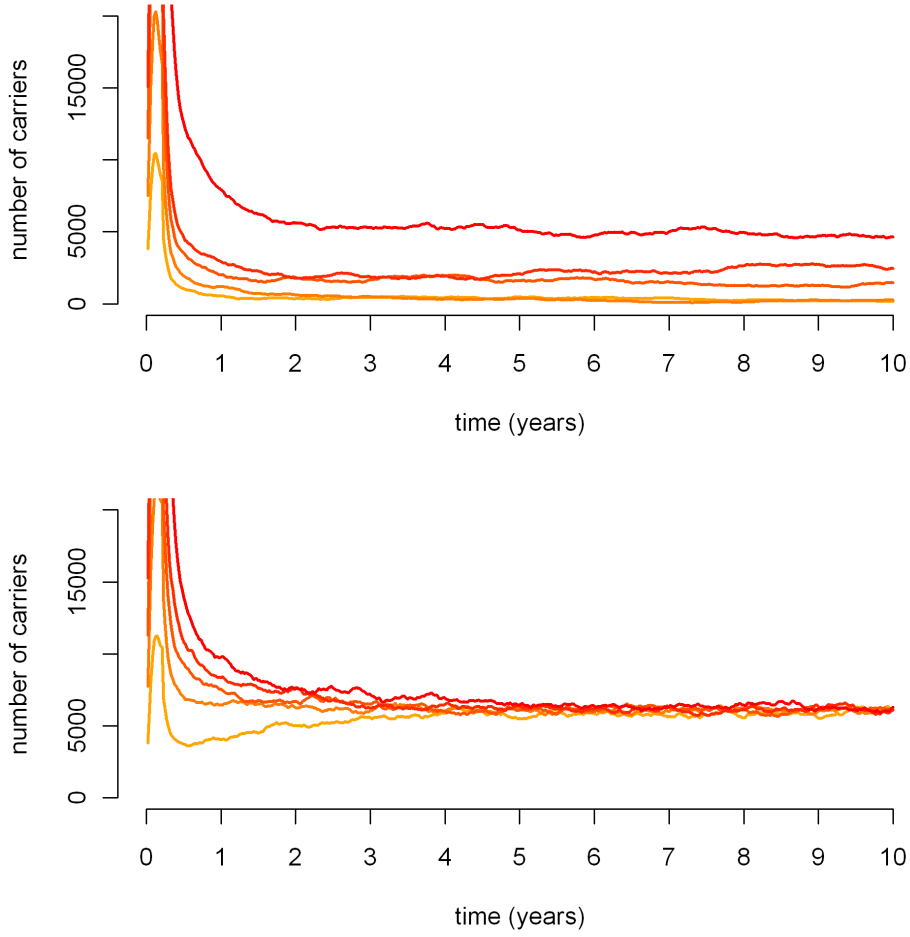


Figure 6.9: The serotype specific number of carriers after initialisation. The intensity of colours represent the rank order of prevalence at initialisation of the otherwise identically parameterised types. The upper panel shows parametrisation reflecting model 3 and the lower panel model 4. The convergence of all serotypes towards the same prevalence illustrates the violation of the population genetic neutrality for model 4.

6.3.7 Vaccination

Introducing infant vaccination against the two most prevalent serotypes leads to extinction of these types and replacement with non-vaccine types in both the vaccinated and - through herd immunity - in unvaccinated individuals (Figure 6.10). Paradoxically, vaccination against all serotypes - at the same efficacy and coverage does not cause extinction of all types and results in only minor indirect effects. Indeed, under the high force of infection scenario the bivalent vaccine is more effective at reducing adult carriage than the universal vaccine. If this vaccine succeeds the bi-valent one it can cause re-emergence of the previously controlled types. This arises because the vaccine types are controlled by a combination of vaccination effect at reducing their reproduction number (by removing susceptibles) and competition from the other types. When the competition is lessened, through the use of a universal vaccine, these high prevalence (high R_0) types can re-emerge (see figure 6.11).

6.4 Discussion

The analysis of an individual-based and yet structurally simple simulation model including acquisition induced short-term immunity against *S. Pneumoniae* is presented here. Non-specific and specific short-term immunity were found to be capable of governing the patterns of competition and coexistence. Stochastic effects in low prevalent serotypes may result in apparent epidemics, if these serotypes are highly invasive (as has been observed in several countries with serotype 1, for instance). The model suggests that high carriage prevalence observed in developing country settings [20] and in native populations [21] might arise from a less effective immune response (due to malnutrition, genetic differences or other factors) rather than mixing patterns alone (see section 6.3.4). Moreover, the model was utilised to assess the impact of vaccination, and contrast a vaccine targeted at the most prevalent serotypes (approximating PCV7) with a broad-based vaccine (e.g. a protein-based one), with the same level of efficacy and coverage. Paradoxically, the low-valency vaccine can lead to greater indirect effects (herd immunity), if the high prevalence (most fit) serotypes are eliminated by a combination of vaccine protection and competition. This benefit over the universal vaccine would be amplified in invasive disease if the types included in the low-valency vaccine are of high invasiveness (as is the case for PCV7). Releasing the competition,

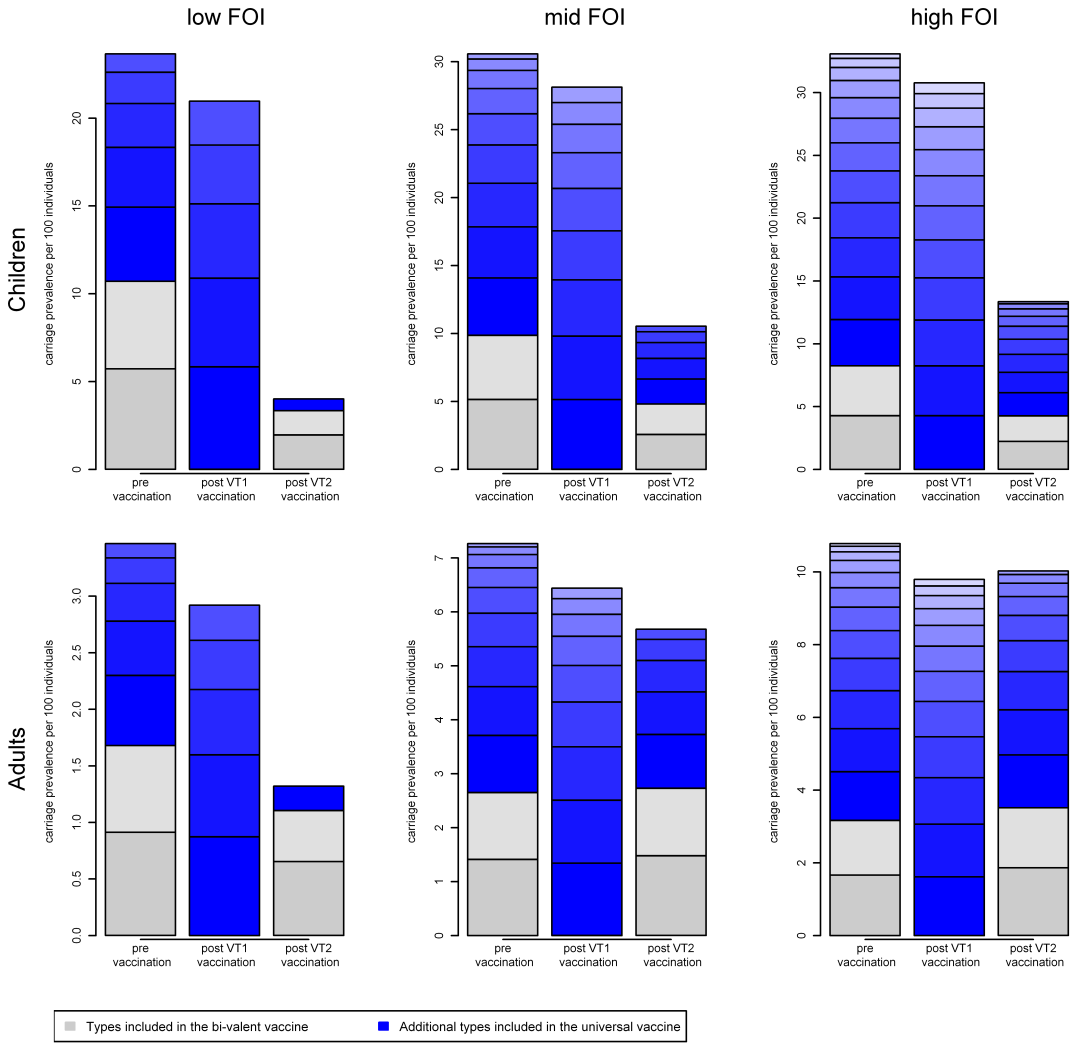


Figure 6.10: Equilibrium mean prevalence of different serotypes by age group (<10 years of age = children, others = adults) under different vaccination programmes: a bi-valent vaccine targeting the 2 most prevalent types (VT1), and an universal vaccine (VT2) with the same efficacy per type. The duration of non-specific and specific immunity was 9 and 18 weeks respectively.

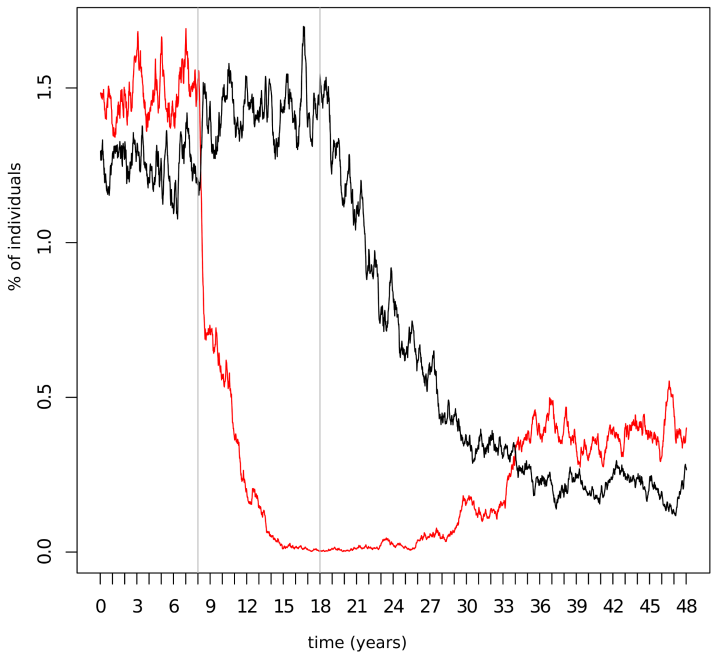


Figure 6.11: Re-emergence of previously controlled serotypes. Bi-valent vaccination was introduced at year 8 after initialisation and the universal vaccine was introduced 20 years after initialisation. The prevalence of all serotypes except two is not displayed for the purpose of clarity. The fittest serotype (in red) is included in the bi-valent vaccine and is controlled shortly after vaccination. With introduction of the universal vaccine below the vaccination threshold the prevalence of all other serotypes declines slowly (the most fit of the serotypes not included in the bi-valent vaccine is shown in black) and as a results reduces the amount of competition which allows the previously controlled type to re-emerge. For this graph the duration of specific and non specific were 18 and 9 weeks respectively in the low FOI scenario.

by using a broad valency vaccine below the vaccination threshold for elimination of carriage, can further lead to a rebound in the previously controlled types, if reintroduced.

The biological basis of protection against colonisation and clearance of carriage of *S. pneumoniae* is poorly understood. From the patterns of pneumococcal acquisition, clearance and re-infection found in longitudinal carriage studies, the existence of short term serotype-specific and non-specific protection can be inferred [10, 11]. With these two mechanisms in the model the specific ecological and epidemiological patterns of the pneumococcus could be generated. Short-term specific and non-specific protection might arise from a combination of both the innate and the adaptive immune response. While an activated innate response, which is associated with increased macrophage or other myelomonocytic cellular activity, might limit initial growth of the pneumococcus in the mucosa, the innate response is short lived and unlikely to be the main effector for clearance of pneumococcal colonisation which may persist for prolonged periods (up to 5 months) before clearance [18]. However, the innate response probably has the role of activating an adaptive immune response that develops to both the serotype-specific capsule and the conserved surface proteins that produce cross-reactive antibodies. These responses could be mediated by locally produced antibody from mucosal B-cells which persist after the pneumococcus has been cleared and prevent initial growth, or enhance clearance in the event of a subsequent heterologous exposure (via non specific antibody) or homologous exposure (both non specific and serotype specific antibody). Although essential in murine models, the role of T-cells in clearance in humans is, as yet, unclear [22, 23].

No mechanism was included in the model to track colonisation events over an individual's lifetime to infer long-term immunity which could be mediated by both capsule specific and non-specific antigens. It was decided to only include its observed effects, a serotype non-specific decline in both carriage duration and the probability of transmission given contact by age. Including long-term immunity and a general maturation of the immune system would complicate parameterisation but is unlikely to affect the mechanism of competition and coexistence. However, it does have an effect of the predicted outcome of vaccination. If vaccination limits colonisation in the vaccinated children the lack of exposure in those early years will leave the the older population more vulnerable. Hence not accounting for exposure driven protection reflects a best case assumption for vaccination effects.

Most scenarios in this chapter are presented setting the duration of non-specific immunity to 9 weeks. This was chosen on the basis of multiple carriage being found to be very low using current standard detection methods [14, 24–27]. However recently developed and yet to be established methods yield higher rates of multiple carriage [28]. If one accounts for that one would need to decrease the duration of non-specific immunity in the simulations which then would lead to an even greater ease for serotype coexistence.

In the absence of further knowledge assumptions had to be made on several aspects of the model:

- Newborns were not protected from colonisation by maternal antibodies. However, this is thought to have little impact on the general transmission dynamics studied here, because very young children are considered to not contribute significantly to the overall transmission dynamics.

- Multiple carriers have the same probability of transmitting each strain as a single carrier has. While this is likely to overestimate the FOI the parameter space which is assumed to be of main interest only allows for few multiple carriers and therefore its impact is believed to be rather minor.

- Non-specific immunity is modelled to allow instant protection since activation of the innate immune response is thought to be rather quick. However, allowing for a slight delay until activation would result in an increased time span for multiple acquisition and reduced competition.

- Different fitness of the pneumococcal serotypes in the model is determined solely by their duration of carriage. Weinberger et al. found that the capsule size of the pneumococcal serotypes is correlated with both avoidance of neutrophil-mediated killing and prevalence [16] but the existence of other factors influencing a serotype's fitness cannot be excluded.

- The non-specific and specific immunity in the model are thought to reflect short term immunity after acquisition. Longer term immunity from multiple infections over an individual's lifetime is hard to distinguish from general maturation of the immune system and is therefore introduced in the model through an age dependent declining duration of carriage and a declining probability of infection given an infectious contact.

Other models have studied competition between serotypes typically employing

competition as a factor limiting the FOI for additional acquisition once infected in a deterministic two strain model [2,3,5]. The model developed here was shown to be flexible and correspond closely to many of the published models when parameterised accordingly (see section 6.2.4). While providing a useful tool to estimate the degree of competition and therefore inferring the possible effects of vaccination on the dynamics of vaccine versus non-vaccine types [4] questions about model validity in the absence of an explicit underlying mechanism have been raised for published deterministic 2-type models [6,7]. It is shown here that the most reasonable parameterisations for the IBM, where the duration of specific immunity exceeds the duration of non specific immunity, provide a close match to the deterministic approach most frequently used for modelling competition of two types and that serotype specific immunity may be the main factor violating the concept of model neutrality. Zhang et al. identified direct competition through physical presence or activated innate immune response as the more likely source of competition over cross-reacting antibodies [29]. In the simulations a similar mechanism could be identified; while the non-specific immunity is the determinant for competition longer lasting specific immunity allows for coexistence of multiple serotypes.

Specific immunity in the simulation is reflected through an increased duration of immunity against acquisition of infection with a previously carried serotype. A different approach towards modelling serotype specific immunity would be that knock-off effects for the homologous serotype are reduced as done in Colijn et al. [7]. With this approach the inclusion of serotype specific immunity would not result in a violation of ecological neutrality. However, the mechanisms of specific and non-specific immunity in the model presented here are biologically plausible and therefore reflect an existing biological mechanism rather than an intrinsic modelling artefact in the context of competing pneumococcal serotypes. Therefore the deterministic model of serotype competition (model 4) as introduced by Lipsitch et al. [3], which is implicitly assuming full serotype specific protection for the course of carriage, should be appropriate for describing the serotype dynamics in the simplified case of only two different pneumococcal serotypes although it formally violates model neutrality.

Cobey and Lipsitch recently proposed a slightly different model to address the means of serotype specific and serotype non specific immunity for the contribution to serotype competition and coexistence [30]. Differences include the representation of both specific and unspecific immunity and homologous co-infection.

Serotype specific immunity is assumed to follow the first clearance and to be imperfect but lifelong. The protective effect from serotype non specific immunity is assumed to both reduce acquisition rates during carriage and reduce carriage duration of all serotypes based on previous exposure. With increasing exposure the duration of carriage becomes increasingly similar for the various serotypes which further act to support coexistence. Homologous acquisition is possible while being infected and the acquired serotype is treated as a separate copy. There is insufficient knowledge of the transmission dynamics and immunity against nasopharyngeal carriage to support either their modelling approach over the one developed in this chapter another, or vice versa. However, the structural findings are similar. Cobey and Lipsitch find that serotype specific immunity reduces the fitness advantages and act to stabilise competition. Acquired serotype non specific immunity further reduces fitness differences. Together specific and non specific immunity patterns can reproduce coexistence of multiple pneumococci despite their competition and differences in fitness.

Sixty-four newborns were chosen to enter the population each week which determines the population size to be 243,792 on average. A higher population size in the model reduces the variance in prevalence and so, in a scenario without importation of cases, this slightly reduces the chance for serotypes to become extinct. Population size further affects the variance of the standardised prevalence. While the qualitative behaviour does not change (long term volatility of low prevalence serotypes) the variance is generally smaller. However, even with a simulated population of 4 million individuals, peaks with a preceding 2-3 times increase in prevalence were common. Spatial diversity and a network structure of infectious contacts would further increase variability and have not been considered in the model.

Long-term temporal trends have been observed in infrequently carried serotypes, in particular serotype 1 [31–33], and made interpretation of the epidemiology of the pneumococcus challenging (see section 3.1.1 and [4]). In the simulation model two potential mechanisms for this occurrence were observed: long duration of serotype specific immunity cause an outbreak-like behaviour. However, the peaks are found to be relatively frequent and the effect is more pronounced in more frequently carried types, which stands in contrast to observed data. Another potential cause of this epidemiologic occurrence was identified: infrequently carried serotypes are most prone to high variance of the standardised prevalence due to stochastic effects of the pathogen population where the apparent peaks

can spread over several years. If a serotype has high invasive potential this effect would be amplified in surveillance systems which monitor invasive pneumococcal disease. This possibly should be kept in mind when trying to interpret these temporal changes.

Vaccination in the model framework with a vaccine formulation consisting of the two most prevalent serotypes (10% of all types) and an efficacy of 65% led to extinction of the targeted types in all scenarios tested. Yet the effect on overall carriage was marginal due to serotype replacement by the untargeted types. No significant overall decline has been observed in most nasopharyngeal carriage studies assessing the impact of vaccination with the 7-valent pneumococcal conjugate vaccine [24, 26] except one in the Netherlands [25]. This apparent difference to the model could be due to various reasons including the lack of power; e.g. to be 90% certain to detect a significant decline from 42% to 39% prevalence as observed in the under 5 year olds in the presented scenario one would need to test 5625 children in each of the two cross-sectional studies. Such large samples are rarely used. It was also shown that vaccines that rely on the benefit of competition induced by the untargeted types in order to control the targeted ones can pose epidemiological challenges when their valency is extended and competition is substantially reduced, like a slow decline in prevalence of the targeted types or re-emergence of vaccine serotypes.

Little is known about the reproduction number of the pneumococcus. Hence, multiple scenarios were studied assuming at least a theoretical chance for each serotype to persist ($R_0 > 1$). Standard methods for the inference of R_0, as discussed in section 1.4.2, are not applicable here. For example the assumption of lifelong immunity which is needed to estimate R_0 from the average age of infection is violated for the pneumococcus. Also it was shown in section 6.3.4 that R_0 cannot necessarily be inferred by the prevalence of the pneumococci and that this is rather dependent on the duration on immunity patterns. However, longitudinal carriage studies could present one way to create insight into the transmissibility of the pneumococci. In the absence of further knowledge on such the resulting strength of competition and hence the threshold for successfully eliminating transmission by vaccination is uncertain.

The concept of pathogens competing for the same ecological niche and the implications for vaccine programmes might be applicable to other settings too. Meningococcal serogroup B vaccines that appear to elicit a robust immune re-

sponse against a broad range of meningococcal strains are in development [34,35]. Whether these vaccines will affect the prevalence of either the other MenB strains or other meningococcal serogroups remains to be seen. Inter-species competition has also been reported. Besides the apparent supporting role of influenza for pneumococcal disease as discussed in section 2.7 reduced carriage of *Staphylococcus aureus* in individuals colonised with the pneumococcus has been detected suggesting that there may be an increase of *S. aureus* after introduction of the pneumococcal conjugate vaccine [36–38].

A structurally simple simulation model which includes serotype specific and non-specific immunity in order to investigate their potential contribution to the epidemiology of the pneumococcus is presented here. This basic mechanism which is likely to be a combination of the innate and the adaptive immune response is sufficient to grant the key epidemiological features of the pneumococcus: coexistence, competition and multiple carriage. This should be subject to evaluation by studies on immunological markers in the context of longitudinal carriage studies. If proven valid this has further implications for vaccination in a competitive environment, and optimal vaccination strategies which effectively reduce the burden of disease without leading to adverse public health consequences need to be evaluated.

6.5 The impact of competition on vaccination effects

A result of the simulation model which may be potentially important for the consideration of higher valency vaccines stands out. At a similar effectiveness as observed for PCV7 a bi-valent vaccine in the simulation model eliminated transmission of the two targeted serotypes in the population and lead to replacement by the other types. However, a vaccine offering the same effectiveness of protection against all serotypes fails to eliminate transmission and could therefore have less population impact. The reasons for this will be investigated with the help of a deterministic two-strain model in this section.

In section 1.3 the threshold theorem for successful eradication of an endemic pathogen in a simple context (SI model) was introduced. In brief: transmission

needs to be reduced by a proportion of at least

$$1 - 1/R_0$$

in order to control that pathogen. For a model similar to the diamond model discussed in section 4.3.4, i.e. when two pathogens are in competition, Lipsitch showed that this threshold is altered depending on the amount of competition between the two pathogens. The amount of competition (and the fitness) induced by the pathogen which is not included in the vaccine determines the reduced threshold for successful vaccination. Assuming that the vaccine is 100% effective at preventing an individual from transmitting infection with pathogen 1 (e.g. by immunising against infection) the coverage f_c needed to be

$$f_c = 1 - \frac{R_{02}}{R_{01}\left[1 + c_1(R_{02} - 1)\right]}$$

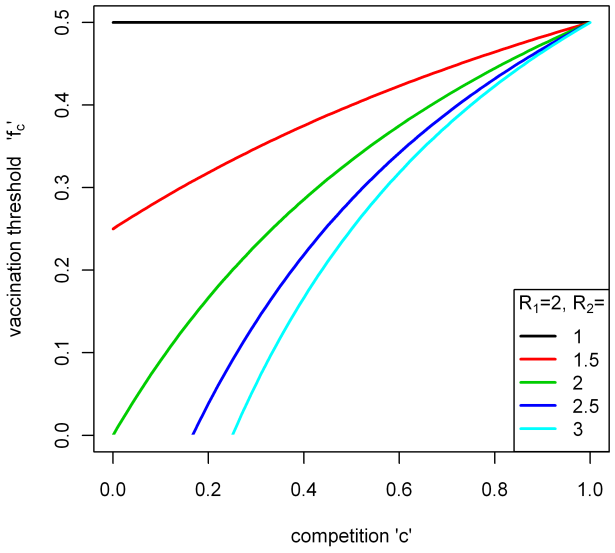


Figure 6.12: Effect of competition and fitness of a second pathogen on the vaccination threshold. The reproduction number for the vaccine type was assumed to be 2.

Here, c_1 is the competition induced on acquisition with type one by carriage with the second type (remember, that $c_1 = 1$ corresponds to no competition and $c_1 = 0$ to competitive exclusion). Figure 6.12 illustrates the effect of competition on the vaccination threshold. If the competing type has a reproduction number smaller or equal to one the type cannot coexist and therefore the vaccination threshold is

not affected (compare sections 1.2 and 4.3.4 for the epidemic threshold theorem). Also when there is no competition induced by the second type, the threshold remains unchanged. However, for competing types with $R_{01}, R_{02} > 1$ the vaccination threshold can be significantly altered in this model and corresponds to the findings of the more complex individual-based modelling approach in this chapter. Figure 6.12 also shows the results on coexistence from section 4.3.4: no vaccination is needed to control type 1 in scenarios where $R_{01} < R_{02}$ and the competition between them is high.

These findings can be interpreted in different ways:

(i) If types are in competition with one another then selective vaccination (preferably against those which are most invasive) can be achieved at a lower coverage or with a less efficient vaccine compared to vaccines targeting all (many) types.

(ii) A combination of vaccine efficacy and coverage found sufficient for a selective vaccine (e.g. PCV7) does not necessarily meet the vaccination threshold for a higher valency vaccine with similar type-specific efficacy and uptake. If implemented re-emergence of some types (as found for the individual-based model) could be the result and the average age of infection would increase (compare section 1.3).

Estimates for the competition (employing the diamond type model framework) between vaccine and non-vaccine serotypes of the pneumococcus following the introduction of the 7 valent vaccine in E& W vary between competitive exclusion and a 53% reduced chance for acquisition while carrying the heterologous serotype depending mainly on the assumption of pre-vaccination trends in IPD to continue [4]. Employing the same techniques based on replacement in IPD observed in the US Melegaro and colleagues found only a 15% reduced probability for acquisition while carrying the heterologous serotype [5]; but this might be underestimating competition because of a change in inpatient blood-testing during the post vaccination era in the US (see section 3.2 and [1]). Estimates obtained from Scottish post vaccination data in section 5.2 data yielded a plausible range from no competition to a possible 55% reduction in acquisition rates. One needs to keep in mind here that all these estimates quantify the competition introduced by the vaccine serotypes rather than the one for the non-vaccine serotypes, which would be important to measure the vaccination threshold. However, these estimates are very susceptible to variations in post vaccination surveillance which

could be the main reason for this wide range of parameter estimates. The arguably best estimates come from a study of longitudinal carriage in Danish children. Auranen et al. found that rates of acquisition were reduced by more than 90% in carriers when compared to non-carriers [2]. However, for the estimation of competition from pneumococcal carriage data the lack of detection of multiple serotypes (compare section 2.5.3) could have introduced an unmasking bias (this is not the case for IPD since invasive disease is believed to result from a single serotype).

No estimate of the reproduction number R_0 for the pneumococci has been reported in the literature to date, so $R_{01} = R_{02} = 2$ is assumed here as an example. That would mean that a perfectly efficient vaccine with a population coverage in order to reduce the transmission by

$$f_c = 1 - \frac{1}{R_{01}} = 1 - \frac{1}{R_{02}} = 50\%$$

would be needed to control both types (e.g. types included and not included in PCV7) at the same time. Assuming that acquisition rates of carriers are reduced by 90% ($c_1 = c_2 = 0.1$), the control of either one type would only require a reduction in transmission of

$$f_c = 1 - \frac{R_{02}}{R_{01}\left[1 + c_1(R_{02} - 1)\right]} = 1 - \frac{R_{01}}{R_{02}\left[1 + c_2(R_{01} - 1)\right]} = 9.1\%!$$

and even $c = 0.5$ would only yield the need for a 33.3% reduction in transmission. This emphasizes the findings from the individual-based model, that selective vaccination strategies could be more effective than approaches targeting as many types as possible.

Summary

In this chapter the dynamics governing the ecology of *S. pneumoniae*, i.e. competition and coexistence, are explored. Based on evidence from longitudinal carriage studies, the means of both short-term specific and non-specific immunity were employed in a parsimonious individual-based model. For a maximum of 20 serotypes the transmission dynamics based solely on acquisition of infection from infected individuals and the thereby induced immunity were studied. It turned out that short-term non-specific immunity acts to reduce the available

pool of susceptibles and hence introduces a competitive interaction of the types. Short-term type-specific immunity on the other hand limits a single types spread and thereby can regulate the coexistence of multiple serotypes. This does provide a biologically plausible mechanism for the observed inter-dependencies of the pneumococci.

Introduction of vaccination against a limited amount of serotypes in this framework yielded similar indirect effects as observed after introduction of PCV7, i.e. herd immunity and serotype replacement. Further, the effects of selective vaccination were compared to those of vaccines with high valency both by simulation of the individual-based model and analytically in the deterministic diamond type model from section 4.3.4. The competition induced by types not targeted by vaccination was found to significantly help with controlling of the vaccine types. If a vaccine is targeting all types, the benefit of competition is lost and the vaccine has to be more efficient or applied at a higher coverage to make up for this.

This marks the end of original research of the thesis. In the next chapter an overview of the main findings will be given and the perspective for future research in this area in order to improve the public health impact of pneumococcal vaccination will be set.

Bibliography

[1] Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease after pneumococcal vaccination. Lancet 6736: 1–12.

[2] Auranen K, Mehtala J, Tanskanen A, Kaltoft MS (2010) Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. American journal of epidemiology 171: 169–176.

[3] Lipsitch M (1997) Vaccination against colonizing bacteria with multiple serotypes. Proceedings of the National Academy of Sciences 94: 6571–6576.

[4] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[5] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, et al. (2010) Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMC infectious diseases 10: 90.

[6] Lipsitch M, Colijn C, Cohen T, Hanage W, Fraser C (2009) No coexistence for free: neutral null models for multistrain pathogens. Epidemics 1: 2.

[7] Colijn C, Cohen T, Fraser C, Hanage W, Goldstein E, et al. (2010) What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae*? Journal of the Royal Society, Interface / the Royal Society 7: 905–19.

[8] Darwin C (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. London: John

Murray, 80–130 pp.

[9] Melegaro A, Choi Y, Pebody R, Gay N (2007) Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. American journal of epidemiology 166: 228–35.

[10] Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, et al. (2008) Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. Journal of Infectious Diseases 197: 1511–8.

[11] Granat SM, Ollgren J, Herva E, Mia Z, Auranen K, et al. (2009) Epidemiological evidence for serotype-independent acquired immunity to pneumococcal carriage. Journal of Infectious Diseases 200: 99–106.

[12] Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine 5: e74.

[13] Rinta-Kokko H, Dagan R, Givon-Lavi N, Auranen K (2009) Estimation of vaccine efficacy against acquisition of pneumococcal carriage. Vaccine 27: 3831–7.

[14] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[15] Melegaro A, Gay NJ, Medley GF (2004) Estimating the transmission parameters of pneumococcal carriage in households. Epidemiol Infect 132: 433–441.

[16] Weinberger DM, Trzccinski K, Lu YJ, Bogaert D, Brandes A, et al. (2009) Pneumococcal capsular polysaccharide structure predicts serotype prevalence. PLoS Pathogens 5: e1000476.

[17] Tsay RS (2005) Analysis of Financial Time Series. 2nd edition, 35–39 pp.

[18] Sleeman KL, Griffiths D, Shackley F, Diggle L, Gupta S, et al. (2006) Capsular serotype-specific attack rates and duration of carriage of *Streptococcus pneumoniae* in a population of children. Journal of Infectious Diseases 194: 682–688.

[19] Turner P, Hinds J, Turner C, Jankhot A, Gould K, et al. (2011) Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. Journal of clinical microbiology .

[20] Hill PC, Cheung YB, Akisanya A, Sankareh K, Lahai G, et al. (2008) Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clinical Infectious Diseases 46: 807–14.

[21] Millar EV, O'Brien KL, Zell ER, Bronsdon MA, Reid R, et al. (2009) Nasopharyngeal carriage of *Streptococcus pneumoniae* in Navajo and White Mountain Apache children before the introduction of pneumococcal conjugate vaccine. The Pediatric Infectious Disease Journal 28: 711–6.

[22] Lu YJ, Gross J, Bogaert D, Finn A, Bagrade L, et al. (2008) Interleukin-17A mediates acquired immunity to pneumococcal colonization. PLoS Pathogens 4: e1000159.

[23] Moffitt KL, Malley R (2011) Next generation pneumococcal vaccines. Current opinion in immunology 23: 407–13.

[24] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[25] Spijkerman J, van Gils EJM, Veenhoven RH, Hak E, Yzerman EPF, et al. (2011) Carriage of *Streptococcus pneumoniae* 3 Years after Start of Vaccination Program, the Netherlands. Emerging infectious diseases 17: 584–591.

[26] Huang SS, Hinrichsen VL, Stevenson AE, Rifas-Shiman SL, Kleinman K, et al. (2009) Continued impact of pneumococcal conjugate vaccine on carriage in young children. Pediatrics 124: e1–11.

[27] Darboe MK, Fulford AJ, Secka O, Prentice AM (2010) The dynamics of nasopharyngeal *Streptococcus pneumoniae* carriage among rural Gambian mother-infant pairs. BMC infectious diseases 10: 195.

[28] Turner P, Hinds J, Gould K, Turner C, Jankhot A, et al. (2010) Conventional techniques for detecting nasopharyngeal pneumococcal carriage significantly underestimate the prevalence of multiple serotype carriage. Tel Aviv, p. 27.

[29] Zhang Y, Auranen K, Eichner M (2004) The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. Epidemiology and infection 132: 1073–81.

[30] Cobey S, Lipsitch M (2012) Niche and Neutral Effects of Acquired Immunity Permit Coexistence of Pneumococcal Serotypes. Science 335(6074): 1376–80.

[31] Miller E, Andrews NJ, Waight PA, Slack MP, George RC (2011) Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. The Lancet Infectious Diseases 11: 760–768.

[32] Black S (2010) The Volatile Nature of Pneumococcal Serotype Epidemiology: Potential for Misinterpretation. The Pediatric Infectious Disease Journal 29: 301–303.

[33] Harboe ZB, Benfield TL, Valentiner-Branth P, Hjuler T, Lambertsen L, et al. (2010) Temporal trends in invasive pneumococcal disease and pneumococcal serotypes over 7 decades. Clinical Infectious Diseases 50: 329–337.

[34] Giuliani MM, Adu-Bobie J, Comanducci M, Aricò B, Savino S, et al. (2006) A universal vaccine for serogroup B meningococcus. Proceedings of the National Academy of Sciences of the United States of America 103: 10834–9.

[35] Jiang HQ, Hoiseth SK, Harris SL, McNeil LK, Zhu D, et al. (2010) Broad vaccine coverage predicted for a bivalent recombinant factor H binding protein based vaccine to prevent serogroup B meningococcal disease. Vaccine 28: 6086–93.

[36] Bogaert D, van Belkum A, Sluijter M, Luijendijk A, de Groot R, et al. (2004) Colonisation by *Streptococcus pneumoniae* and Staphylococcus aureus in healthy children. Lancet 363: 1871–2.

[37] Regev-Yochay G, Dagan R, Raz M, Carmeli Y, Shainberg B, et al. (2004) Association between carriage of *Streptococcus pneumoniae* and Staphylococcus aureus in Children. The Journal of the Amercian Medical Association 292: 716–20.

[38] Regev-Yochay G, Lipsitch M, Basset A, Rubinstein E, Dagan R, et al. (2009) The pneumococcal pilus predicts the absence of Staphylococcus aureus cocolonization in pneumococcal carriers. Clinical infectious diseases 48: 760–3.

7 Conclusions

7.1 Summary and main findings

The introduction of pneumococcal conjugate vaccines into childhood vaccination programmes has substantially reduced the burden of pneumococcal disease in England & Wales, Scotland and elsewhere. Due to herd immunity effects vaccination has not only provided protection for vaccinated individuals but for unvaccinated individuals as well. However, currently available vaccines only provide protection against a very limited amount of the more than 90 different capsular serotypes. These serotypes vary greatly in their resistance to clearance by the immune defence mechanisms and their ability to invade the bloodstream once they have colonised the nasopharynx. Furthermore, the serotypes, despite their stable coexistence, compete for the same ecological niche. This generates an important feature of pneumococcal ecology: serotype replacement following selective conjugate vaccination. The extent of the serotype replacement and the pathogenicity of the replacing serotypes greatly impacts on the overall population effects of vaccination. This thesis had the aim of analysing the effects of pneumococcal conjugate vaccination on both disease and carriage and studying modelling concepts regarding their suitability for an adequate description of pneumococcal ecology patterns and vaccine induced changes to it.

When estimating the impact of vaccination an assumption on the baseline, the number of cases if the vaccine had not been introduced, has to be made. Most studies assume this baseline to be similar to the number of cases one or two years before the start of vaccination. Upward trends in IPD incidence data in England & Wales prior to vaccination were found [1]. Hence, the assumption of an unchanged baseline several years after the introduction of PCV7 was questionable and a comparison to control pathogens, which are similarly dependent on blood

culturing practice but have not been subject to vaccination, was undertaken (section 3.1.1 and [2]). Generally increasing incidence amongst these pathogens prior to introduction of PCV7 was found, perhaps indicating increasing ascertainment. However, the size of the increase wasn't consistent amongst the pathogens for all age groups and the trends did not continue after introduction of the pneumococcal vaccine in all cases. However, the assumption of either a constant baseline after vaccination or a continuous pre vaccination trend in the post vaccination era provided lower and upper bounds for the observed incidence of the control pathogens and hence could be understood as a range of likely baselines. These results highlight the general value of good pre vaccination data to detect ongoing changes which might bias the analyses of the impact of vaccination. In particular, an estimate for additional uncertainty in the baseline arising from pre vaccination data was obtained, which assisted with the analysis of the impact of PCV7 on IPD in England and Wales [3] and similar studies can help with the analysis of vaccine impact in various settings.

No analysis of post PCV7 epidemiology in Scotland has been available. Increasing incidence of IPD prior to the introduction of conjugate vaccination was found similarly to England and Wales. Hence, both methods, assuming no post vaccination trend in ascertainment and assuming the pre vaccination trend to continue, were employed for the analysis (section 3.1.2). Disregarding the different methodologies PCV7 vaccination was shown to have significantly reduced the invasive disease burden associated with the targeted serotypes in both vaccinated and unvaccinated individuals in Scotland. While there was little evidence for serotype replacement in the age groups up to 64 years, an increase in non vaccine serotype IPD balanced out the decrease amongst vaccine types in the elderly. These findings contribute to a better understanding of the impact of pneumococcal vaccination in Scotland. The analysis shows the general success of the 7-valent conjugate vaccine in reducing the overall burden of IPD in the Scottish population, particularly in children. Also it highlights the need for a better understanding of the underlying changes in pneumococcal carriage in Scotland in order to better understand the reasons for the differences in serotype replacement amongst the age groups, i.e. whether serotype replacement in nasopharyngeal carriage was largely confined to the elderly as observed in IPD or whether the different, probably more pathogenic, serotypes were associated with serotype replacement amongst the elderly population.

Vaccine induced changes in the predecessor of IPD, nasopharyngeal carriage, are

key to the understanding of the ecological basis for any changes in pneumococcal disease. These changes were assessed from studies conducted both before and after introduction of PCV7 in healthy children and their family members in Hertfordshire, England (see section 3.1.3 and [4, 5]). Rates of vaccine type carriage declined amongst all age groups while non-vaccine type carriage increased. In contrast to invasive disease, where an overall decrease was found amongst all age groups [3], serotype replacement was complete in carriage. This suggests that the pathogenicity of the replacing non-vaccine types is lower than that of the vaccine types. Similarly, the additional serotypes in the 10 and 13 valent vaccines were estimated to be highly pathogenic As a result vaccination with either of them is likely to result in a further reduction of IPD even if serotype replacement in nasopharyngeal carriage is assumed to be complete. This provides a simple method how the combination of invasive disease data and data on pneumococcal carriage can determine the likely impact of future conjugate vaccination formulations and can help informing vaccine policy making. Furthermore, the observation of full serotype replacement in pneumococcal carriage attests strong competition between the serotypes for the same ecological niche and the general importance for pneumococcal ecology in the impact of conjugate vaccination.

An important aspect about pneumococcal ecology is the coexistence of pneumococcal serotypes despite their competition and an apparent fitness advantage of those types which were found to resist clearance from the nasopharynx for longer. The reasons for this feature of pneumococcal ecology are unknown. In chapter 4 basic dynamic deterministic modelling approaches for two competing serotype groups are reviewed and examined for their ability to permit serotype coexistence. The model structure of which stable coexistence is a common outcome is not neutral [6], i.e. coexistence is artificially promoted by the model structure rather then induced by explicit assumptions on the transmission dynamics. However, the model structure does implicitly assume immunity against homologous reinfection for the duration of carriage while it allows heterologous reinfection at a reduced rate. This serotype specific immunity could limit the spread of one type, thereby promoting the spread of the other and thus support coexistence.

Following an approach of Choi et al. [7] a model which permits serotype coexistence as a common outcome was employed (see section 5.2). A Markov Chain Monte Carlo approach was taken to evaluate its fit to Scottish IPD data. The model provided a good fit to overall vaccine type and non-vaccine type IPD incidence in Scotland. However, when stratified into age groups the model showed

inconsistencies with the dynamics in some of the age groups. That is, the model could not replicate little serotype replacement in children and extensive replacement in the elderly population at the same time. While PCV7 successfully reduced the burden of disease in Scotland this shows that the underlying dynamics are not understood well enough to re-model the observed vaccine effects on invasive disease. Hence further studies are needed in order to make the outcome of future pneumococcal vaccination more predictable. With the existence of pneumococcal carriage data one could study whether the discrepancy between the model and the data arises from differences in pathogenicity of the replacing serotypes in the elderly population and other age groups or indicates a lack of understanding of pneumococcal transmission dynamics, and hence an invalid model structure, or more likely arises from higher stochasticity of the data than anticipated.

In order to investigate the possible role of short-term homologous immunity against reinfection in the coexistence of pathogens, in particular the pneumococci, a parsimonious individual-based model was developed. It comprises 20 types which differ in their resistance to clearance and hence duration of carriage. Types compete through infection induced immunity against heterologous acquisition. It is found that the number of serotypes which stably coexist is associated with the duration of homologous immunity, i.e. without type specific immunity coexistence of multiple types is impossible. The model further shows a strong dependence of carriage prevalence on the duration of immunity which suggests that high pneumococcal carriage rates amongst native populations and some developing countries could be mainly due to an inferior host defence mechanism (e.g. an impaired immune response possibly because of malnutrition) rather than increased transmission due to extensive population mixing. The model also served to illustrate a general feature about vaccination in a competitive environment: selective vaccination benefits from the competition induced by the non-vaccine types and hence reduces the vaccination threshold for interrupting transmission of those types. These findings have important implications for the consideration of higher valency vaccines. Higher valent or universal vaccines have a higher vaccination threshold. In other words, while the effectiveness and uptake of PCV7 were sufficient to suppress transmission of vaccine serotype in the population this might not necessarily be the case for vaccines of significantly higher valency which could lead to failure to provide herd immunity to unvaccinated groups at high risk for disease.

7.2 Future work

The 13 valent pneumococcal conjugate vaccine has now replaced PCV7 in the UK's national immunisation program. First results suggest a further reduction in IPD cases in England & Wales [8]. Vaccines of higher valency are in development. This thesis has given rise to a number of questions which can be critical for an evaluation of the likely impact of higher valency pneumococcal vaccines. Pneumococcal conjugate vaccines are being rolled out across the developing world, where rates of carriage and disease are considerably higher than in the developed world. Understanding the impact that these vaccines may have and the best strategy to mitigate future risks of replacement effects is critical for public health.

Cross sectional pneumococcal carriage studies carried out in Scotland in children and the elderly population would be helpful to determine the source of the pronounced replacement amongst IPD in the elderly population despite little evidence for replacement in other age groups. One could determine the replacement effects amongst pneumococcal carriage in children and the elderly population. If contrary to the observation in IPD the amount of serotype replacement in nasopharyngeal carriage was similar across the age groups one could investigate the differences between the specific replacement serotypes to look for potential differences in invasiveness. The elderly population has been observed to have the highest rates of invasive pneumococcal disease incidence, hence determining the source for pronounced serotype replacement in this age group could help adjusting future vaccination strategies to achieve higher population benefits from pneumococcal vaccination in Scotland. However, after the introduction of PCV13 in Scotland cross sectional carriage studies for neither the prevaccination baseline nor the changes after introduction of PCV7 can be carried out anymore. Based on the assumption that the invasiveness of a serotype is similar in England & Wales and in Scotland and similar for all age groups one could infer approximate carriage rates in Scottish children and elderly from the combination of IPD data in Scotland and case:carrier rates from E& W.

With the introduction of PCV13 only a few years after the introduction of PCV7 changes in pneumococcal ecology associated with PCV7 are likely to be still ongoing (see chapter 3.1.3). Hence, the incremental impact of vaccination with the extended valency will be impossible to determine directly from the data directly via relative risks. The diamond-type model discussed in sections 4.3.4 and 5.2

and its expansion to include vaccination with PCV13 (section 5.3.5) could be fitted to emerging data from the post PCV13 era in England & Wales. Also the results of the post PCV7 carriage study and those of an upcoming post PCV13 carriage study could be included in the fitting process for a better determination of replacement in pneumococcal carriage and the calculation of the case:carrier ratios. The results could give both an evaluation of the incremental impact of PCV13 and a prediction of its long term effects to help better predicting the public health benefits of pneumococcal conjugate vaccination in E& W. Further this work would provide a baseline for the prediction of the impact of vaccination with higher valency vaccines.

A major caveat of all analyses determining aspects of the transmission dynamics of the pneumococci is the lack of detection of multiple carriage (see section 2.5.2). With newly developed laboratory methods it becomes clear that the current gold standard method for detection of the pneumococcus from carriage isolates significantly underestimates the amount of multiple carriage [9]. However, the relative abundance of serotypes is usually skewed towards the dominance of one serotype. It is yet unclear to what extend multiple carriage affects the transmission of the pneumococcus from a host which is infected by multiple types. Information on this would be critical to distinguish the validity of different model structures and their assumptions on pneumococcal transmission. Nasopharyngeal colonisation data collected trough a longitudinal study design in households could provide insight on how multiple carriage affects the transmission dynamics and the role of abundance of a serotype. Transmission events between household members could be analysed for the abundance of colonisation of the transmitted type in the source case. This work would be essential for the understanding of the transmission dynamics of *S. pneumoniae* and would provide a basis for future predictions of the population impact of pneumococcal conjugate vaccination.

In longitudinal studies pneumococcal carriage was found to reduce both homologous and heterologous acquisition rates. The analysis in chapter 6 shows that this heterologous immunity might induce the competition between the serotypes while an extended span of homologous immunity permits serotype coexistence despite the fitness advantages of some types. A similar study by Cobey and Lipsitch [10] finds that serotype specific immunity, which is imperfect but lifelong, can also permit coexistence. There is insufficient evidence on natural immunity against pneumococcal colonisation to favour either of the two approaches. Correlates for natural immunity against pneumococcal carriage are poorly understood

which makes a measure of the extent and duration of serotype specific immunity challenging. Highly powered longitudinal carriage studies which monitor the effects of previous homologous carriage on acquisition rates for multiple months could be one way to achieve this. These studies could provide further insight if and how protection against homologous acquisition influences the ecology of the pneumococcus. With that knowledge the accuracy of the prediction of the impact of future pneumococcal vaccines could be vastly improved.

The individual based model developed in chapter 6 was built to gain general insight about the possible role of short-term immunity against acquisition of carriage in pneumococcal ecology. The model can be extended to fit other purposes. A household and spatial structure and can be added to the population to study if these less random mixing patterns are likely to further contribute to serotype coexistence in the population. Also the number of serotypes reflected by the model can be increased in order to parameterise the model according to the the specific serotypes of the pneumococcus and the layer of invasiveness can be added to infer changes in pneumococcal disease from predicted carriage rates in the model. The parameterisation could be assisted with the fitting to longitudinal carriage data which would allow estimation of the force of infection and hence the competitive pressure which drives serotype replacement and determines the vaccination threshold. Pneumococal conjugate vaccines with increasing valency are being developed. Furthermore, vaccines targeting common proteins might even prove effective against most of the pneumococci. While serotype specific effectiveness estimates yield important information about the direct effects of the new vaccines, mathematical models will be needed in order to determine if the new vaccines are likely to sufficiently limit the spread of the pneumococci to provide herd immunity. With good serotype specific data it should also be possible to parameterise the individual based model accordingly to determine an optimal vaccine design which makes use of the competitiveness of the pneumococci, that is to work with the natural ecology of the organism for the benefit of public health. In order to study the effects of childhood vaccination in adults one may also need to include an exposure dependent acquisition of serotype non specific immunity. This could lead to increased carriage rates in the adult population if vaccination fails to eliminate transmission of the targeted serotypes.

The research presented in this thesis is focussed on *Streptococcus pneumoniae* but could help with the analysis of other pathogens as well. *Nesseria meningtis* has a similar polysaccharide capsule and when conjugated to a protein becomes

a powerful immunogen. Carriage rates of serogroup Y are high and the recently licensed quadrivalent A,C,Y,W135 vaccine might cause replacement with other, more virulent, meningococal serogrups [11]. Although structurally different to *S. pneumoniae* vaccine introduction against a limited amount of strains of rotavirus and human papillomavirus (HPV) poses the potential risk for introducing ecological changes which may lead to unanticipated indirect effects.

Bibliography

[1] Trotter CL, Waight P, Andrews NJ, Slack M, Efstratiou A, et al. (2010) Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. Journal of Infection 60: 200–208.

[2] Flasche S, Slack M, Miller E (2011) Long term trends introduce a potential bias when evaluating the impact of the pneumococcal conjugate vaccination programme in England and Wales. Eurosurveillance 16: 1–6.

[3] Miller E, Andrews NJ, Waight PA, Slack MP, George RC (2011) Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. The Lancet Infectious Diseases 11: 760–768.

[4] Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. Epidemiol Infect 133: 891–898.

[5] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in England: A Cross-Sectional Study. PLoS Medicine 8: e1001017.

[6] Lipsitch M, Colijn C, Cohen T, Hanage W, Fraser C (2009) No coexistence for free: neutral null models for multistrain pathogens. Epidemics 1: 2.

[7] Choi YH, Jit M, Gay N, Andrews N, Waight P, et al. (2011) 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? PLoS ONE 6: e26190.

[8] Health Protection Agency (2010). Pneumococcal disease. URL

http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Pneumococcal/.

[9] Turner P, Hinds J, Turner C, Jankhot A, Gould K, et al. (2011) Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. Journal of clinical microbiology .

[10] Cobey S, Lipsitch M (2012) Niche and Neutral Effects of Acquired Immunity Permit Coexistence of Pneumococcal Serotypes. Science 335(6074): 1376–80.

[11] Trotter CL, Greenwood BM (2007) Meningococcal carriage in the African meningitis belt. The Lancet infectious diseases 7: 797–803.