Imperial College London

Department of Infectious Disease Epidemiology

Mathematical models to evaluate the impact of increasing serotype coverage in pneumococcal conjugate vaccines

Alessandra Løchen

Supervised by

Professor Sir Roy M. Anderson and Dr Nicholas J. Croucher

Submitted in part fulfilment of the requirements
for the degree of Doctor of Philosophy at
Imperial College London, October 2021

I declare that the work presented here is my own, and was carried out under the supervision of Professor Sir Roy M. Anderson and Dr. Nicholas J. Croucher. Any information derived from the work of others has been appropriately acknowledged or cited, and can be found in the list of references.

Alessandra Løchen (2021)

Abstract

Of over 100 serotypes of Streptococcus pneumoniae, only 7 were included in the first pneumococcal conjugate vaccine (PCV). While PCV reduced the disease incidence, in part because of a herd immunity effect, a replacement effect was observed whereby disease was increasingly caused by serotypes not included in the vaccine. Dynamic transmission models can account for these effects to describe post-vaccination scenarios, whereas economic evaluations can enable decision-makers to compare vaccines of increasing valency for implementation. This thesis has four aims. First, to explore the limitations and assumptions of published pneumococcal models and the implications for future vaccine formulation and policy. Second, to conduct a trend analysis assembling all the available evidence for serotype replacement in Europe, North America and Australia to characterise invasive pneumococcal disease (IPD) caused by vaccine-type (VT) and non-vaccine-types (NVT) serotypes. The motivation behind this is to assess the patterns of relative abundance in IPD cases pre- and post-vaccination, to examine country-level differences in relation to the vaccines employed over time since introduction, and to assess the growth of the replacement serotypes in comparison with the serotypes targeted by the vaccine. The third aim is to use a Bayesian framework to estimate serotype-specific invasiveness, i.e. the rate of invasive disease given carriage. This is useful for dynamic transmission modelling, as transmission is through carriage but a majority of serotype-specific pneumococcal data lies in active disease surveillance. This is also helpful to address whether serotype replacement reflects serotypes that are more invasive or whether serotypes in a specific location are equally more invasive than in other locations. Finally, the last aim of this thesis is to estimate the epidemiological and economic impact of increasing serotype coverage in PCVs using a dynamic transmission model. Together, the results highlight that though there are key parameter uncertainties that merit further exploration, divergence in serotype replacement and inconsistencies in invasiveness on a country-level may make a universal PCV suboptimal.

Acknowledgements

Contents

8

List of Tables

List of Figures

12

13

15

Nomenclature

AMC Advanced Market Commitment

AOM Acute otitis media

BAC Bacteraemia

BF   Bayes Factor

CAP Community-acquired pneumonia

CCR Case-to-carrier ratio

CEA Cost-effective analysis

CFR Case-fatality ratio

CI   Confidence interval

CrI   Credible interval

CUA Cost-utility analysis

DALY Disability-adjusted life years

DTM Dynamic transmission model

EE   Economic evaluation

EMP Empyema

FOI   Force of infection

GAVI Global Alliance for Vaccines and Immunization

HIC   High-income country

hPCV Hypothetical pneumococcal conjugate vaccine

ICER Incremental cost-effectiveness ratio

IPD   Invasive pneumococcal disease

IRR   Incidence rate ratio

LMIC Low- or middle-income country

LYG   Life years gained

LY   Life years

MCMC Monte Carlo Markov Chain

MEN Meningitis

MLST Multi-locus sequence typing

NIPD Non-invasive pneumococcal disease

NPNM Pneumococcal non-pneumonia, non-meningitis

NP   Nasopharyngeal

NTHi Non-typeable *Haemophilus influenza*

NVT Non-vaccine-type serotype

PCV Pneumococcal conjugate vaccine

PNE Pneumonia

PPSV Pneumococcal polysaccharide vaccine

PPV Protein-purified vaccine

QALY Quality-adjusted life years

SEI   Susceptible-exposed-infected model

SEP   Septicaemia

SIR   Susceptible-infected-recovered model

SIRS Susceptible-infected-recovered-susceptible model

SI   Susceptible-infected model

SIS   Susceptible-infected-susceptible model

ST   Sequence type

TD   T cell dependent

TI   T cell independent

VE   Vaccine efficacy

VT   Vaccine-type serotype

WCV Whole-cell vaccine

Chapter 1

Introduction

The aim of this chapter is to provide some essential background on *S. pneumoniae* regarding the pathology, detection methods of pneumococcal infection, epidemiology as well as treatments and control by various interventions to reduce disease induced by pneumococcal infection. The chapter will conclude with an overview of the content of the thesis and the major aims of the research.

1.1   Background

The bacterium *Streptococcus pneumoniae*, or simply "pneumococcus", was first discovered and isolated in 1881 by Louis Pasteur [13] (France) and George Sternberg [14] (USA) independently. Although previously observed [15–17], it was Pasteur and Sternberg who discovered its virulence by injecting human saliva into a rabbit and observing the bacteria in the rabbit's blood. In 1974, it was renamed from *Diplococcus pneumoniae* (i.e. pairs of cocci), to its present-day name *Streptococcus pneumoniae*, for its resemblance to streptococci (Figure 1.1).

Figure 1.1: Illustration of *Streptococcus pneumoniae*. Reproduced from [5].

1.2   Pathology

The pneumococcus is a Gram-positive bacterial pathogen that grow in pairs and are commensal and opportunistic. It generally colonises the upper respiratory tract of humans, detectable by nasopharyngeal (NP) swabs. The bacteria are transmitted through direct respiratory secretions (i.e. mucus or saliva) in person to person contact [18] and can colonise a person for weeks to months, dependent on their age and serotype [19]. It is asymptomatic in the

nasopharynx but can progress to cause a great proportion of otitis media cases [20–22].

1.2.1 Key surface antigens

Pneumococcus has three major surface layers: the plasma membrane, the cell wall, and the capsule (Figure 1.2). The cell wall, which anchors the capsule, consists of proteins and cell wall polysaccharide. This polysaccharide is a teichoic acid containing phosphoryl-choline, which binds host receptors [23] and allows for the attachment of bacterial choline-binding proteins to the cell surface [24]. Phosphoryl-choline is recognised by pentraxins, C-reactive proteins (CRP) and natural immunoglobulin M (IgM), which are both inhibited by the capsule, thereby allowing pneumococcus to evade both complement activation and, as a result, phagocytosis [25]. The cell wall is found in all pneumococci, whereas the capsule is serotype-specific. The capsule is the thickest and outermost layer that encapsulates the entire bacteria, making it a physical barrier inhibiting the binding of high affinity antibodies to subcapsular proteins [25]. It determines the serotype [26], something that was observed by the reaction of serotype-specific antisera to the capsule [27, 28]. There are around 100 different serotypes of pneumococcus presently identified [29]. The capsule manipulates the complement pathway, resisting phagocytosis [25], and affecting the amount of complement components deposited on its surface [30, 31], ultimately evading the immune system [32]. It plays a key role in virulence and pathogenicity as unencapsulated bacteria were shown to be considerably less pathogenic than their encapsulated counterparts [27, 33]. Non-encapsulated pneumococci (i.e. non-typeable, NT) make up around 3% to 19% of asymptomatic carriage isolates with increasing reports in non-invasive disease and occasional cases in invasive disease patients, potentially due to the selection pressure from vaccines [34].

Of the surface proteins secreted by pneumococci, the most notable are major autolysin (LytA), pneumococcal surface protein C (PspC), pneumococcal surface antigen A (PsaA), and pneumococcal surface protein A (PspA), which show promise as future vaccine candidates [35]. LytA loosens and remodels the cell wall, which is necessary for growth and division, and results in the release of pneumolysin (Ply), required to maintain long-term asymptomatic car-

riage [36], as well as cell wall products [37], and thought to be the key in shedding to promote transmission [38]. PspC, an adhesin, binds to certain receptors to disrupt the host immune function, and can eventually enable invasion into the blood-brain barrier/cerebrospinal fluid [39]. PsaA is a manganese transporter and an adhesin to host cells [40–42], but also serves to resist the host innate immune response's oxidative stress [43]. PspA inhibits complement activation [44] by inhibiting binding of components that would lead to opsonization.
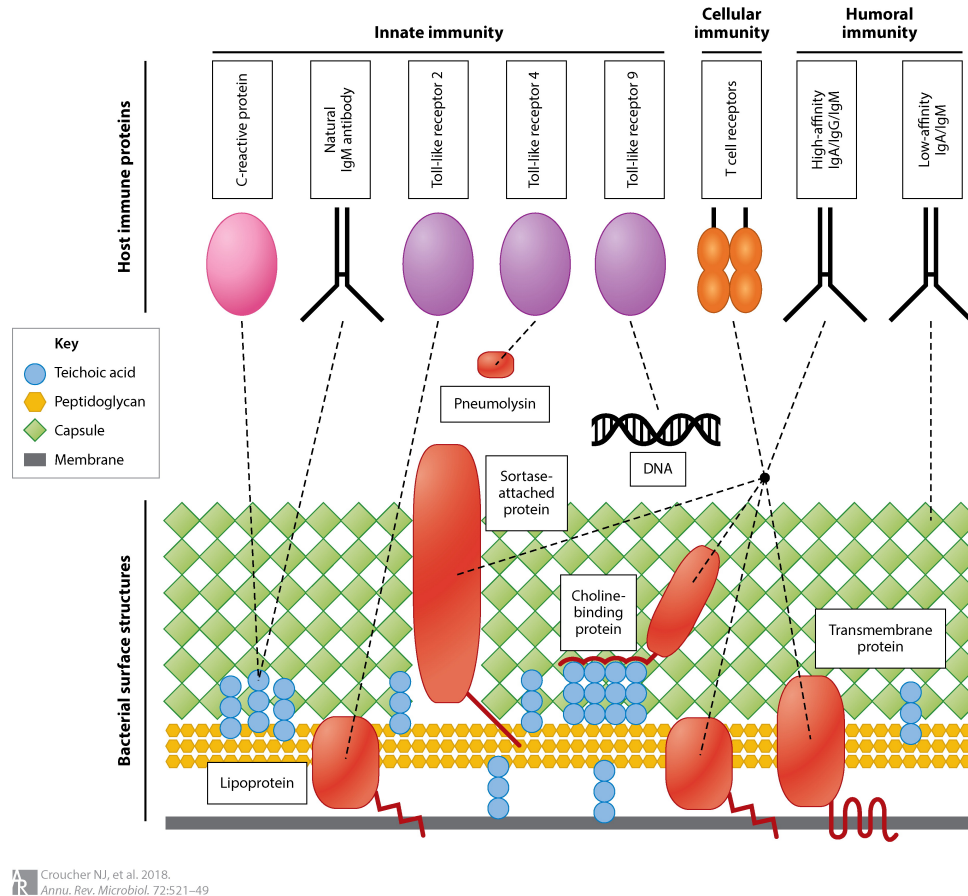
Figure 1.2: Summary of the major interactions between pneumococci and their host. Structures forming the pneumococcal cell surface are shown at the bottom, and host immune recognition proteins at the top. Reproduced from [6].

1.2.2 Immune responses

Colonisation and disease are immunising events that can be mediated by anti-capsular (serotype-specific) antibody immunity for some serotypes [45, 46]. The host's immune response involves two different pathways. First, there is the T-cell independent pathway

(TI), in which the host reacts to the capsule molecule. This triggers the alternative comple-ment pathway [47], which ultimately results in B cells proliferating and differentiating into plasmablasts [48]. These secrete low affinity antibodies (IgM and immunoglobin A; IgA) [48], which recognise a range of pneumococcal capsule polysaccharides, and as capsules are serotype-specific, this pathway is therefore referred to as specific immunity. The low affinity antibodies also recognise adhesins, degradative enzymes and cell wall synthases [24], which are non-specific to serotypes. In the unvaccinated population in the United States, there was an observed decline in carriage and IPD rates for all serotypes that was concurrent with barely any increase in anti-capsular antibodies during the same (pre-vaccination) period [49]. This indicated that anti-capsular immunity was not the primary mechanism of protection against disease, but that one or more other mechanisms may exist, which could be in the form of non-specific immunity targeting proteins or other structures [49].

This non-specific immunity may be innate immune mechanisms (i.e. macrophages, dendritic cells and neutrophils enabling phagocytosis), or T helper 17 (Th17) cells which can recognize pneumococcus in the nasopharynx. There, Th17 cells secrete IL17A, which mobilizes phago-cytic immune cells to clear the bacteria in a non-specific manner [50]. Th17 cells can then form memory T cells [51]. As such, the non-specific immunity mechanisms may be responsible for the age-associated decline in carriage and disease observed in unvaccinated populations [49].

The other pathway is the T-cell dependent pathway (TD), in which the host reacts to pneumo-coccal proteins. Once encountered, these proteins are displayed by antigen-presenting cells to B cells which then recruit CD4+ T cells. These T cells can differentiate into two notable forms. T follicular helper (Tfh) cells trigger activation and differentiation of B cells [52]. B cells either become plasmablasts and secrete low affinity antibodies [48], as in the TI pathway, or they proliferate in germinal centres. There, they can undergo somatic hypermutation and isotype switching, and thereafter secrete high affinity antibodies (IgG and IgA) [48, 52, 53]. These B cells can then either become plasma cells secreting high affinity antibodies or memory cells.

The interplay between specific and non-specific immunity [46, 54–56] increases serotype

diversity while maintaining their coexistence [57]. Although the consequences of these immune responses have not yet been fully defined, particularly for the progression from carriage to disease with or without immunity, immunodeficient individuals are particularly susceptible to IPD [58, 59], indicating that immunity exists and is important. Even though the body may build an immune response to infection, there is limited evidence for host memory induced by NP carriage [60, 61]. While it was shown in one study that consecutive colonisations, regardless of serotype, did not diminish the duration or re-acquisition in infants, and that they could therefore be colonised repeatedly with multiple strains [55], other studies showed that there was evidence that duration of carriage reduced with host age [62, 63] and some serotypes elicited anti-capsular immunity against carriage in Israeli toddlers [46].

1.2.3 Pneumococcal disease

Carriage is the precursor to disease, however, it is still unclear at what rate carriage progresses to disease. This is made more complex because each serotype has a different invasive potential, i.e. odds of invasive disease following colonisation compared to a reference serotype or all other serotypes observed, that is potentially age-dependent [64–66]. The range of serotypes in NP colonisation and IPD varies by geography, age, and socioeconomic conditions [67]. Some serotypes are the main cause of disease outbreaks but are rarely detected in carriage due to short carriage episodes of these newly acquired serotypes or density of colonisation in the nasopharynx [68, 69] and are therefore difficult to predict through colonisation patterns [70]. For example, serotypes 1 and 5 are associated with more severe childhood pneumonia and disease outbreaks in crowded adult settings, but are rarely detected in NP carriage swabs [69], and are therefore considered to be epidemic serotypes. Others, considered the pediatric serotypes, such as serotypes 6B, 19F and 23F, are frequent in carriage but are not common in disease. This is unsurprising given that serotypes differ in their abilities to evade and elicit immune responses, and it has also been suggested that some serotypes are associated with certain clinical presentations (such as serotype 1 with severe pneumonia) [69].

Additionally, while there are some similarities in IPD serotypes globally, serotypes 4 and 18C were more common in Europe and North America pre-vaccination [10] whereas serotypes 1 and 5 were more prevalent in Africa and Asia [10, 71, 72], demonstrating the geographic heterogeneity between serotypes in disease. In carriage isolates for example, serotype 13 was more prevalent in India, Indonesia and Kenya, whereas this serotype was not common at all in European and North American settings [67].

Pneumococcal disease is categorised as either non-invasive or invasive (Figure 1.3). Non-invasive pneumococcal disease (NIPD) will manifest as acute otitis media (AOM), sinusitis, or non-bacteraemic pneumonia (or community-acquired pneumonia, CAP). AOM occurs primarily in children, likely due to their underdeveloped and therefore dysfunctional Eustachian-tube, which is the port of entry for pathogens in the nasopharynx [73]. Invasive pneumococcal disease (IPD) occurs when the bacteria are found in otherwise sterile locations, manifesting as bacteraemic pneumonia (i.e. pneumonia leading to bacteraemia), bacteraemia, or meningitis [74, 75].
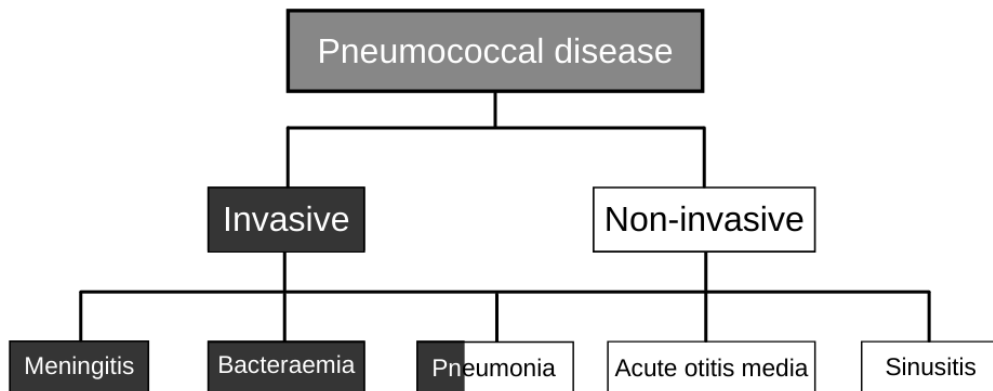
Figure 1.3: Pneumococcal disease classification as invasive or non-invasive. Reproduced from [7].

1.3   Diagnostic methods

Identification of pneumococci is typically culture-based, employing agar plating, optochin susceptibility and bile solubility tests, although non-culture based identification methods using polymerase chain reaction (PCR) have been increasingly developed in the last two decades [1]. Antigen-based tests such as the serotype-specific urine antigen testing can also be used as a rapid diagnostic immunoassay test based on immunochromatography detecting the C-polysaccharide in urine and other body fluids. While it has been suggested to have high sensitivity (77% - 88%) and specificity (67% - 100%) [76, 77], it has been noted to exhibit false positivity especially in the paediatric population [78, 79], cross-reactivity with other streptococci, and prolonged positivity after disease has resolved [80].

Once identified as pneumococci, isolates are serotyped. Current detection methods of pneumococcal serotypes in the nasopharynx include traditional agar culture plating, colony blot assay, PCR, red fluorescent protein (RFP), microarray, as well as quantitative or real-time PCR (qPCR) (Table 1.1). Because of the low sensitivity to multi-serotype carriage among others, the shortfalls of these methods have been pointed out [81–83]. Agar plating, for example, may identify only the most abundant serotype, and additional serotypes with less than a certain frequency will go undetected making it extremely time-consuming and cost-ineffective for adequate detection of multiple carriage [1]. This made it difficult to distinguish between post-vaccination serotype replacement and unmasking, in which only the most dominant serotype of a multi-serotype carrier is detected and therefore minor serotypes become more widely reported after the removal of dominant serotypes [84] and therefore confounded how serotypes truly interacted with one another. The discordance between NP colonisation and IPD serotype distribution observed in some studies may be due to the low sensitivity of serotype-specific detection methods which may not detect minor serotypes in carriage [70, 85].

The Quellung reaction, considered the golden standard for serotyping, is widely used and involves a broth medium inoculated with NP swabs such that pneumococcal serotypes can be detected using a capsular reaction test using a panel of diagnostic antisera [86]. This is

labour intensive for detecting multiple carriage however, because of the lower frequency of minor serotypes [82].

Latex agglutination, a modification of the Quellung reaction, mixes cell suspensions with antibody-coated latex particles and has been shown to be most effective for sweep serotyping in less developed countries as it does not require plating. Similar to the Quellung reaction though, it may not detect low abundance serotypes [81].

PCR, qPCR and real-time PCR have been used extensively with various targets including the autolysin gene (*lytA*) and an iron acquisition adenosine triphosphate-binding cassette (ABC) transporter gene (*piaB*), which are both specific to pneumococcus but not always ubiquitous [87], a limitation of this method. Its potential specificity to other types of streptococci as well as lack of specificity to pneumococci by way of absence of target genes in certain pneumococcal isolates, highlights the need to use multiple targets [88, 89].

On the other hand, microarray serotyping, which involves fluorescently labelling DNA samples and hybridizing them to a microarray, is likely the gold standard for detection of multiple carriage particularly of serotypes at low abundance, but is more expensive [2, 81]. This method was made possible due to the sequencing of the capsular polysaccharide synthesis (*cps*) locus [90].

The dot blot or immunoblot assay, a modification of the Western Blot assay, involves a membrane previously incubated with a monoclonal antibody identifying pneumococcal isolates which can then be re-plated to be probed with serogroup-specific antibodies [91, 92]. This can therefore lack specificity for serotypes or require more effort to distinguish each serotype.

The use of genomic sequencing has been pivotal for defining strains of pneumococci, despite its cost. Multilocus sequence typing (MLST) is a high resolution DNA fingerprint method in which fragments of several housekeeping loci are determined, the combination of which are referred to as a sequence type (ST) [93]. It is therefore highly specific, as it can also be

Table 1.1: Description of common serotyping methods. Adapted from [1] and [2].

Method	Positive predictive value	Limitations	Reference
Quellung reaction	99.6%	Costly; time-consuming; labour intensive; does not detect minor serotypes as well	[1, 2]
Latex agglutination	91.4%	Costly; time-consuming; does not detect minor serotypes as well	[1, 2]
Real-time PCR	89.3%	Does not distinguish closely related serotypes well	[1, 2]
Microarray	93.7%	Costly; requires significant technical expertise for interpretation	[1, 2]
Dot blot	82% [91] - 98.6% [92]	Lack of specificity; requires significant optimisation for each serotype	[1, 2, 91, 92]
Genome sequencing	98% [96] - 99.3% [97]	Costly; requires significant technical expertise	

used to track evolutionary patterns of STs [94]. Whole genome sequencing, i.e. sequencing the entire genetic material, has also been used increasingly, as it has the benefit of providing more granular detail on pneumococcal antigens and their evolution [95].

1.4 Epidemiology

Pneumococcal disease accounted for 8-12% of all deaths in children between 1 and 59 months in 2000 [98], and caused over 1.2 million deaths globally due to pneumococcal pneumonia in 2016 (Figure 1.4, top left) [99]. It was responsible for 73% of deaths due to lower respiratory infections in children less than 5 years and over half the deaths due to lower respiratory infections in people aged over 70 years in 2019 [100], and is the most common cause of bacterial pneumonia [101]. Those most at risk for invasive pneumococcal disease (IPD) are children (especially those under the age of 2), the elderly [102, 103], and the immuno-compromised [104, 105], although adults are also at risk [106], albeit lower (Figure 1.4, bottom row).

While NIPD causes less mortality and is less costly per individual, it is much more frequent than IPD [107]. Prior to the introduction of the pneumococcal conjugate vaccines, pneumococci were the most common cause of AOM [108] and CAP [20–22], with an estimated 30-40% of diagnosed CAP due to pneumococci [109]. Although it is difficult to find a causative agent for pneumonia in many cases due to diagnostic sensitivity unable to distinguish between carriage and disease, prior to vaccination, pneumococci were thought to be responsible for most of the undiagnosed pneumonia [110] and are therefore considered to be one of the most common causes of bacterial pneumonia globally [101].

IPD incidence varies not only by age, but also by geography and socioeconomic status. The incidence of IPD was around 10 fold higher in developing countries than developed countries pre-vaccination [8] (Figure 1.5), although this is also in part due to immunocompromising conditions such as HIV and sickle cell disease. Overall IPD incidence in children less than 6 years before vaccination was between 8 and 24 cases per 100,000 children per year in European countries, between 20 and 56 cases per 100,000 children per year in Australia, Chile, and New Zealand, and between 70 and 75 cases per 100,000 children per year in the US [111]. The higher incidence in the US likely reflects the higher rates of blood culture in outpatient bacteraemia, as incidence of meningitis in the same age group is similar between the US and European countries, which have similar socioeconomic conditions: 3.6 cases per 100,000 people per year in the US vs around 4.6 cases per 100,000 people per year in European countries [111].
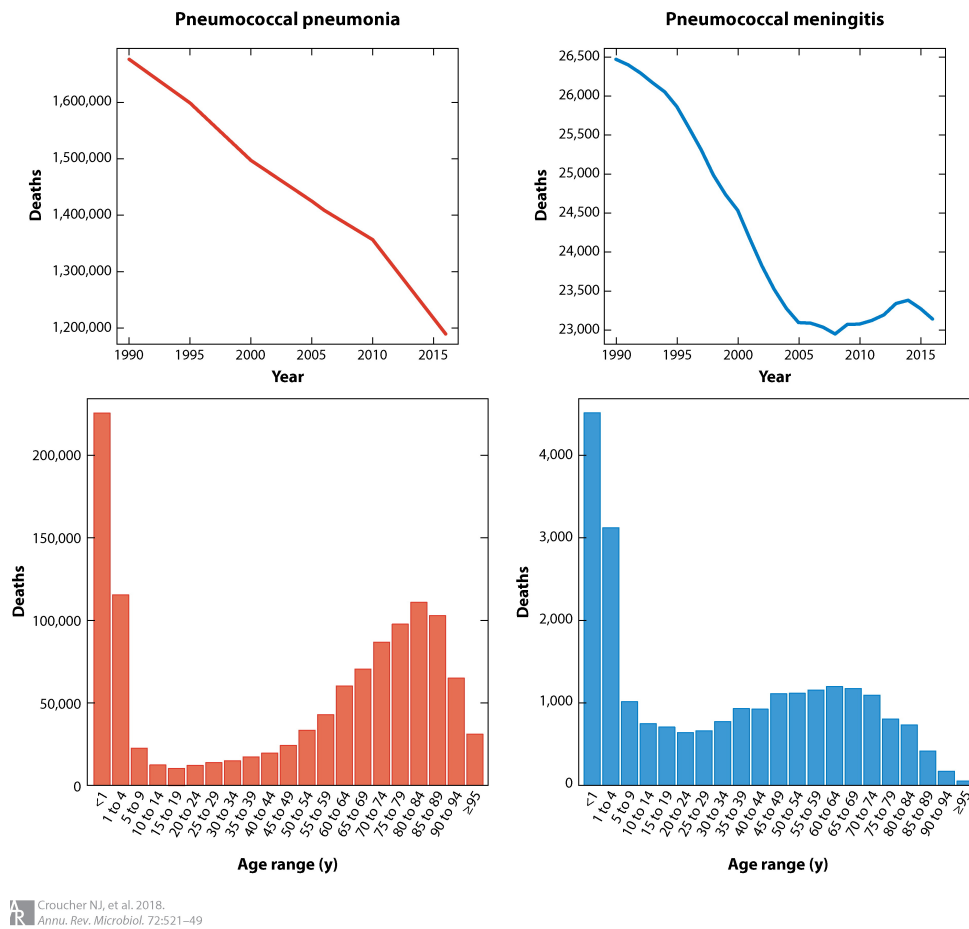
Figure 1.4: Reduction (top) and age distribution (bottom) of global mortality due to pneumococcal pneumonia and pneumococcal meningitis. Reproduced from [6].
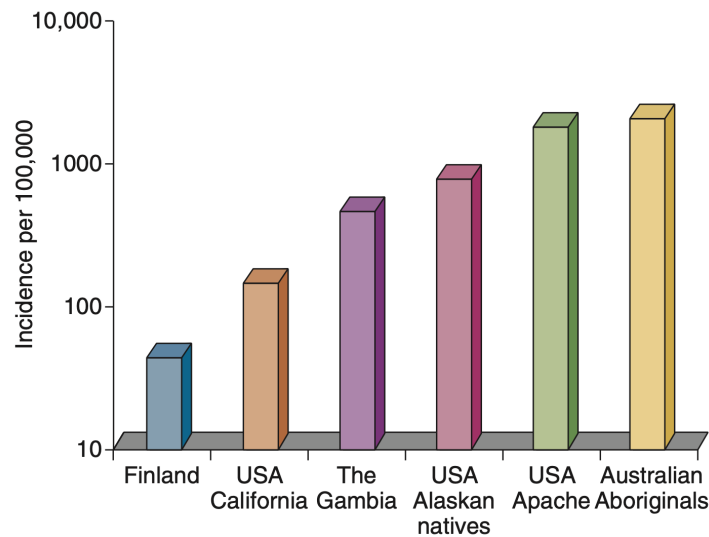
Figure 1.5: Invasive pneumococcal disease incidence in children less than 2 years varies by geographic location and demographic population. Reproduced from [8, 9].

The prevalence of carriage is highest in children under 5 years, where it ranged between 25 and 50% in high income countries (HICs) [112] and 20 and 90% in low- and middle-income countries (LMICs) [113] pre-vaccination. Similarly, the carriage prevalence in adults with no children in the household was less than 10% in HICs [114–116] but between 25 and 50% in LMICs [117–119]. Children are the main source of transmission and are the cause of its introduction into households [19], therefore it is unsurprising that the highest rates of carriage are seen in children attending daycare centres [67, 120]. The carriage prevalence in adults increased with the presence of children at home in a study in the United States, from 6% without to 29% with [114], and carriage rates were increased in school-aged children in Spain with siblings less than two years old compared to those without [121]. Carriage prevalence peaks in children less than 3 years old and steadily declines with age [67].

1.5 Treatments and public health interventions

Antibiotics have been widely used to treat *S. pneumoniae*, particularly AOM, since the use of first sulfanilamide and then penicillin in the early 20th century. When pneumococci developed resistance to these in the late 1970s [122], it resulted in other antibiotics being discovered and employed. These include antibiotics such as erythromycin, penicillin, clindamycin, cephalosporin, rifampicin, vancomycin and trimethoprim-sulfamethoxazole. Pneumococci are competent for transformation and have the ability to pick up DNA for recombination, resulting in the acquisition of resistance genes [123]. The deaths attributable to penicillin- and macrolide-resistant pneumococci increased by 12-40% in European countries between 2007 and 2015 [124]. The degree of antibiotic pressure in a population influences the emergence of newly resistant serotypes, in turn changing the serotype distribution [125, 126]. Using MLST-derived data, the eBURST algorithm was applied to pneumococcal strains and found that antibiotic-resistance appeared to have occurred within a rare allelic profile that then increased greatly under strong selection [94]. Pneumococci are thought to be exposed to antibiotic pressure when in the nasopharynx with prolonged carriage, particularly as this is when they are also exposed to other species that can pass antibiotic resistance genes [69].

Emerging resistant strains and a persistent high mortality rate in vulnerable populations prompted the development of a novel intervention to which immunity could be built and mortality reduced; namely vaccines [127].

Initial pneumococcal vaccine development occurred as early as 1911 with a trial in South African miners with whole-killed pneumococcus [128], followed by a series of trials suggesting adult protection with capsular materials [129–132]. In 1945, this was proven by MacLeod et al when they showed protection conferred by pneumococcal polysaccharides in military barracks [133], and a hexavalent polysaccharide vaccine was thereafter licensed. Due to the widespread use of antibiotics, particularly penicillin, and other treatments such as chemotherapy, interest in vaccine development waned [122]. Increasing antibiotic resistance saw the renewed interest of vaccines, culminating in the introduction of 14-valent polysaccharide vaccines in the late 1970s and then later the 23-valent polysaccharide vaccine in the mid-1980s which included 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. As a result of the multitude of serotypes circulating in most populations, only a fraction of these, based on the frequency and disease burden generated by specific serotypes (Figure 1.8), could be included in vaccines due to manufacturing restrictions. However, these serotypes accounted for 90% of the global IPD burden in all age groups [134–137]. While this vaccine is effective against IPD in adults [138] and generally well-tolerated in children, it is poorly immunogenic in children, and particularly infants, as it does not induce memory cells and the process by which pure polysaccharide antigens elicit a T-cell independent response is not fully developed in the younger age group [139]. Given that infants are the primary target of vaccination, alternative vaccine formulations were required.

Children respond much better to T-cell dependent antigens compared to pure polysaccharide antigens that are T-cell independent [140], prompting the development of a protein conjugated polysaccharide vaccine eliciting a T-cell dependent response which, as mentioned earlier, allows secretion of high affinity antibodies and B cell proliferation, including memory B cells [139]. The development of this pneumococcal polysaccharide-protein conjugate vaccine (PCV) was made possible by the success of the conjugate vaccine against *Haemophilus influenzae*

targeting *H. influenzae* type b which resulted in more than a 98% decline in disease in the US [141].

1.5.1 Pneumococcal vaccines

The first licensed pneumococcal polysaccharide-protein conjugate vaccine (PCV) was a heptavalent vaccine developed against serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F, and conjugated with cross-reactive molecule 197 (CRM_{197}), a mutant of the diphtheria toxin. Since cross-protection between related serotypes occurs, it was hoped that PCV would also provide vaccine-induced cross-protection for serotypes within the serogroups of the serotypes included, requiring only one serotype from each serogroup to be included [136, 137, 142]. However, vaccines typically confer serotype-specific antibody-mediated anti-capsular immunity. Regardless, PCV was proven to be safe, immunogenic and efficacious against IPD in the target population infants and children [140, 143–145]. Vaccine-induced immunity does not reduce carriage duration, but one is less likely to acquire the vaccine-type serotypes (VTs), and thereby block disease progression. Vaccines, as well as antibiotics, have succeeded in reducing the burden of pneumococcal disease in children and the elderly, i.e. the most vulnerable populations.

Before vaccination, global deaths due to pneumococcus were estimated to be between 700,000 and 1 million each year, with around 14.5 million episodes of serious disease [98]. Of the deaths in HIV-negative children, over 60% were estimated to occur in ten African and Asian countries. Six to eleven serotypes accounted for more than 70% of IPD cases worldwide [10]. Combined, these observations highlight the importance of developing an effective vaccine against pneumococcal disease.

The serotypes included in the first PCV were the dominant serotypes in carriage and IPD in infants in the US, and also reflected many of the serotypes that had a high likelihood of antibiotic resistance [146]. This first vaccine iteration excluded serotypes 1, 5 and 6A, which were among the seven most common serotypes causing IPD globally (Figure 1.6) [10].

Serotypes 1 and 5 were among the top 4 most common IPD serotypes in Africa, Asia and Latin America [10]. Clinical trial endpoints for PCV7 were to reduce IPD and NP colonisation. Two clinical trials based in the US aimed to prove reduced IPD: one in healthy Californian infants [145], and the other in American Navajo and White Mountain Apache children less than 2, a historically at-risk indigenous population with high IPD rates relative to the rest of the US population (Figure 1.5) [140]. These showed 97.4% and 76.8% efficacy respectively against IPD caused by vaccine type serotypes (VT). Another clinical trial of PCV9 (PCV7 + 1, 5) conducted in infants in South Africa showed an 83% reduction in VT IPD, and 65% reduction in VT IPD in HIV-infected children [147]. Combined, these clinical trials resulted in a vaccine efficacy against VT of 93% [148]. Importantly, clinical trials also showed a protective effect against VT NP colonisation as well [149–151] and suggested that vaccine-induced immunity reduced VT acquisition [150, 151].

As formulations were based on the serotype landscape in the US, PCV7 was highly effective in North America, achieving 89% reduction in IPD in a prelicensure trial [145], and a 70% reduction in IPD in children under 2 years (the vaccine target population) in a surveillance study [152]. In European countries, which have similar demography and serotype distributions to the United States [10], the vaccine was considered effective [153–156]. It was estimated that between 2000 and 2020, PCVs averted around 610,000 deaths in children under 5 in LMICs [157]. Vaccination reduced IPD in both vaccinated and unvaccinated populations, as the unvaccinated population benefits from herd immunity effects [158–164], evidenced by the fall in IPD in adults where adults over 65 years of age had an 18% reduction in IPD [152]. Between 2000 and 2015, mortality due to pneumococcal disease was reduced globally by around 51%, and serious episodes of pneumococcal disease were reduced by around 75% down to 3.7 million in HIV-uninfected children between 1 and 59 months [11]. In 2010, more than 75% of the deaths in children under five occurred in African and Southeast Asian regions [165], and in 2015, half of the deaths due to pneumococcus occurred in four African and Asian countries: India, Nigeria, Democratic Republic of Congo, and Pakistan (Figure 1.7).
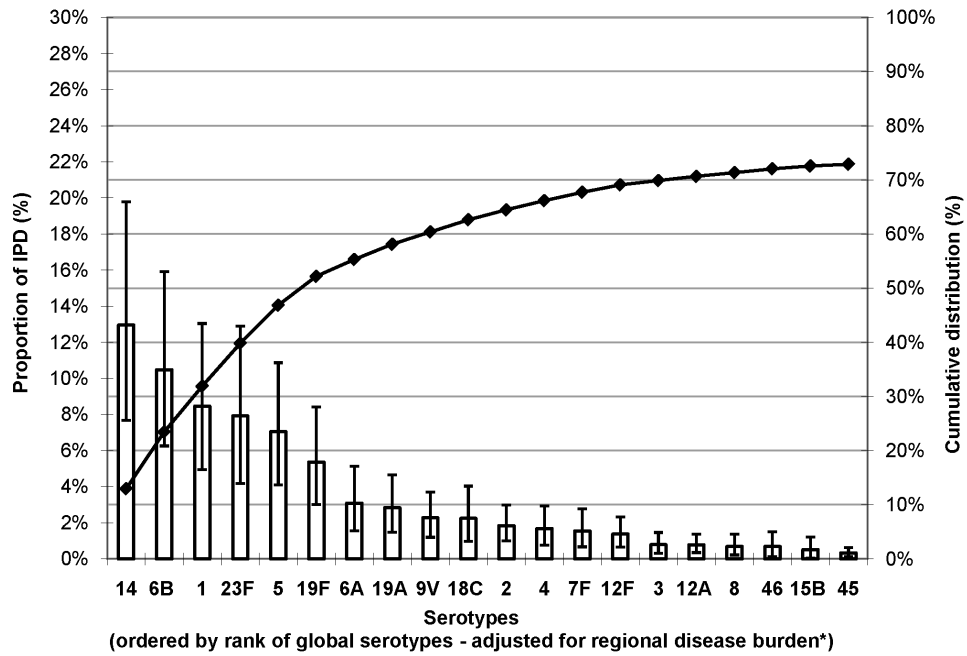
Figure 1.6: Cumulative proportion of disease caused by the top serotypes causing IPD globally in children < 5 years old. Reproduced from [10].
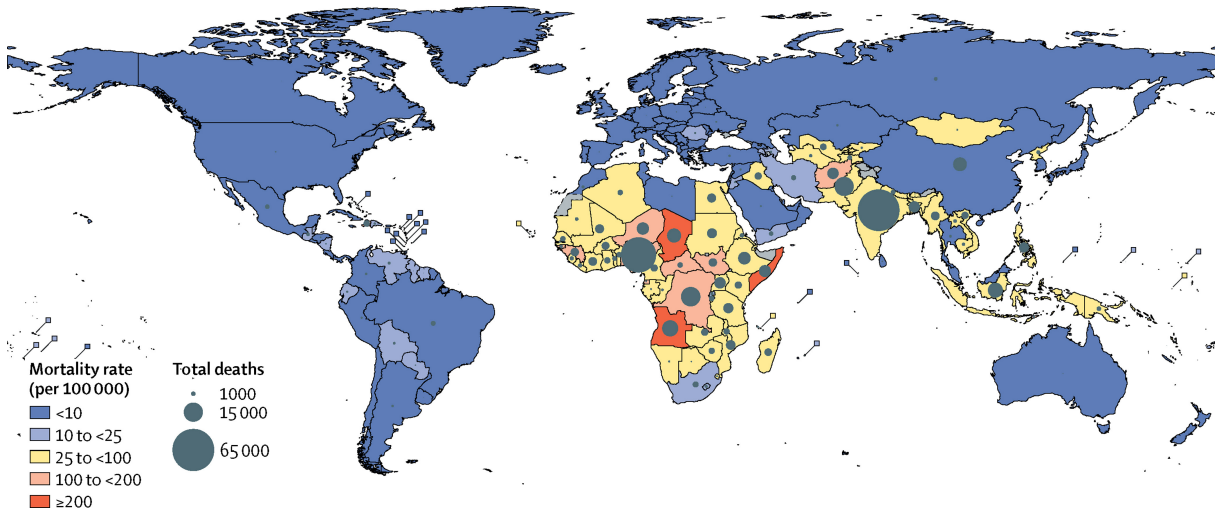
Figure 1.7: Global mortality rates of deaths due to pneumococcus in HIV-uninfected children 1-59 months in 2015. Reproduced from [11].

Serotype replacement

While the vaccine largely reduced IPD incidence in all age groups [166] due to the reduction in VT, a replacement effect was observed where non-vaccine serotypes (NVT) caused an increasing number of IPD cases [167, 168]. Although nontypeable *Haemophilus influenzae*

(NTHi) replaced pneumoccus post-vaccination and became the leading cause of AOM immediately after vaccination with PCV, NVT drove the proportion of AOM caused by these two otopathogens to later be equivalent [108]. As such, serotype replacement might affect AOM more than IPD [169, 170], as AOM is not known to be as serotype-specific. The serotype replacement effect was also observed in carriage as well [112, 171, 172], where the removal of VT colonising the nasopharynx by vaccines likely removed them as competitors to NVT, thus allowing NVT to proliferate in this space [170, 173]. This caused overall pneumococcal carriage prevalence in some cases to be unchanged post-vaccination [174]. Replacement observed in carriage may at least partially also be a result of serotype unmasking, where limitations of diagnostic methods identified only the most dominant serotypes and therefore made the presence of NVT more apparent post-vaccination with the removal of VT [173, 175]. Although both cases are likely to contribute to the replacement observed [175, 176], the pre-vaccination prevalence of NVT multi-carriage would need to be substantially higher to maintain the overall carriage prevalence observed post-vaccination if unmasking was the driver of this phenomenon. However, multi-carriage prevalence was not observed to be higher than the 10-30% range [83, 171] or higher overall pre-vaccination [177], strongly suggesting an increase in NVT prevalence in carriage post-vaccination rather than unmasking [175].

This increase in NVT carriage prevalence may also be caused by serotype switching, a process in which pneumococci acquire a different capsule via genetic transformation and recombination of the *cps* locus. The rate of serotype switching differs between capsules [178]. It occurs primarily within serogroups, likely reflecting the reduced recombination effort required to switch between immunologically-similar capsules [90], and may provide insight into how pneumococcal dynamics will change with future serotype-specific vaccines [178, 179]. The impact of vaccine pressure on serotype switching is still unclear, as serotype switching is an intrinsic feature that predates vaccine introduction [95, 180].

Whole genome sequencing and MLST played a crucial role in identifying genomic changes within the pneumococcal population after vaccination. Pre-vaccination, MLST identified

variations of STs within serotypes, and therefore serotypes were associated with particular strains and clonal complexes [93, 94], most of which tended to be dominated by one or two serotypes. However, this association was eroded post-vaccination by serotype switching, as isolates that were identical at the MLST loci expressed different serotypes [181]. MLST identified that the serotype replacement was due to new strains coming in that had always expressed NVTs, expansions of resident strains that expressed NVTs, and apparent switching of some strains from VTs to NVTs. Whole genome sequencing confirmed this and additionally was able to identify multiple switches happening in parallel within individual strains post-vaccination [182]. Genomic sequencing found that pre-existing switches that were either rare or imported were responsible for the post-vaccine switching events that emerged post-PCV [183]. This rarity explains why switches were not detected pre-PCV, and how changes in selection pressures such as vaccine introduction reveal diversification within lineages [182]. Additionally, genomic studies have shown that negative frequency-dependent selection of certain other loci on pneumococcal strains preserves certain gene clusters present in both VT and NVT and may predict serotype replacement in carriage [184].

Serotype replacement reduces the effectiveness of the vaccines [160, 173] for both vaccinated and unvaccinated populations, who benefit from herd immunity from the vaccinated population. Additionally, PCV7 VT accounted for around 50% of IPD in children less than 5 years globally, and particularly in Africa and Asia [10]. To address this need for wider coverage, higher valent vaccines were developed (Figure 1.8).

Higher valent PCVs

The first expanded vaccine, the 10-valent PCV10 (PCV7 + 1, 5, 7F), is manufactured by GlaxoSmithKline and is conjugated to protein D of NTHi. As NTHi and pneumococcus are the two leading causes of AOM [108] and NTHi causes around one third of AOM cases [185], vaccine developers hoped that PCV10's conjugation to the NTHi protein would elicit a protective effect against NTHi-caused AOM as an added benefit of the vaccine's pneumococcal protection. There is some evidence that PCV10 may protect against NTHi-

related AOM, however this is somewhat contested at present since it seems to be setting-specific [186, 187]. The serotypes in this expanded iteration included epidemic serotypes 1 and 5, as well as serotype 7F which is highly invasive [188]. These serotypes were prevalent in Africa, parts of Asia and parts of South America [189]. With these serotypes included, PCV10 VT accounted for around 70% of IPD in children less than 5 years in Africa and Asia, as well as globally [10].

The other expanded vaccine, the 13-valent PCV13 (PCV10 + 3, 6A, 19A), like PCV7, is manufactured by Wyeth/Pfizer and is conjugated to diphtheria proteins. It also increases the proportion of IPD covered by the vaccine from around 50% to 75% globally and in African and Asian countries. The additional serotypes were chosen because of their increasing incidence in IPD post-vaccination. As well as exhibiting increasing antibiotic resistance [190], serotype 19A became a frequent cause of IPD globally post-vaccination: in the US [191, 192], in Australia [193], in Canada [194], in Europe [167, 195], in Japan [196] and in South America [197]. Serotype 3 is associated with severe disease and increased mortality particularly in adults [188, 189], and serotype 6A, despite some cross-protection with VT 6B [198], is associated with antimicrobial resistance and causes a substantial portion of disease [199, 200].

Both the expanded valency vaccines have shown to be efficacious [189, 201, 202], although pooled trials of these assessed non-inferiority based on correlates of antibody protection due to PCV's success in removing VTs [148]. Based on the Californian, Navajo and South African clinical trials, an antibody concentration of 0.35 μg/ml was used as a threshold for correlate of efficacy against IPD. The percentage of people in which this antibody threshold was met was used as a basis of non-inferiority, although individual antibody response for each serotype was not. In fact, there is evidence of reduced efficacy against serotype 3 by PCV13 resulting in continued prevalence of this serotype in disease [203].
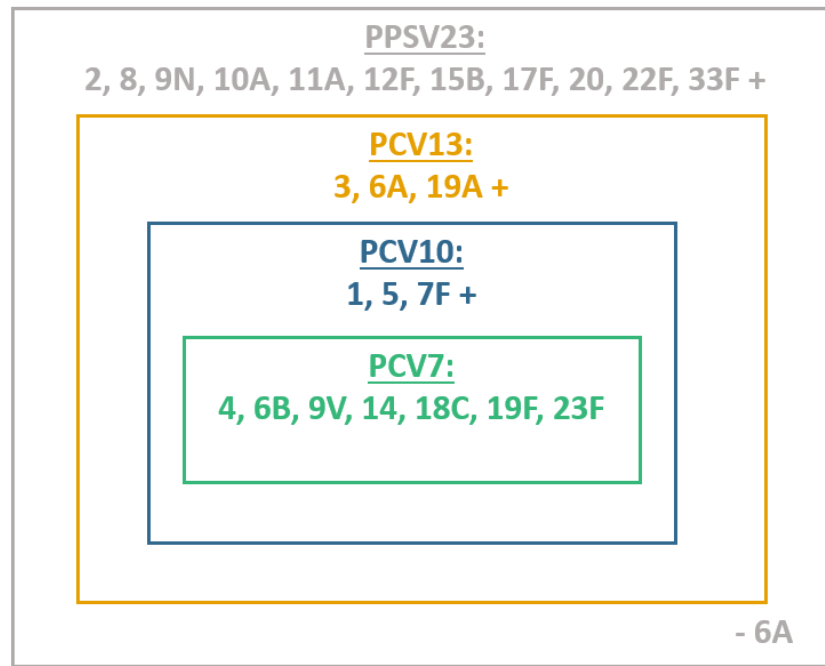
Figure 1.8: Current licensed pneumococcal vaccines are the pneumococcal polysaccharide vaccine with 23 serotypes (PPSV23), and the pneumococcal conjugate vaccines with 7 (PCV7), 10 (PCV10) and 13 (PCV13) serotypes included.

Despite their effectiveness, PCVs are expensive to manufacture because of their protein-conjugated nature and the multitude of pneumococcal serotypes to be added. In 2007, Global Alliance for Vaccines and Immunisations (GAVI) announced a US$1.5 billion donation to sub-sidize pneumococcal vaccines in LMICs. This pneumococcal Advanced Market Commitment (AMC) is a legally binding 10-year commitment from LMICs to purchase the vaccines at a fixed low price (US$3.50/dose) [204]. Both PCV10 and PCV13 fit the target product profile for this, as they were technologically close targets in late-stage clinical trials that covered serotypes in LMICs. Eligibility of countries was based on a minimum 70% coverage of the diphtheria pertussis tetanus vaccine and a maximum income threshold [204]. As of 2021, 152 countries have introduced PCV, of which 60 have received support from the vaccine funder GAVI [205].

Further attempts are being made to target more IPD-causing NVTs and to anticipate which NVTs may emerge and dominate serotype distribution profiles. PCV15 (PCV13 + 22F, 33F), by Merck, and PCV20 (PCV13 + 8, 10A, 11A, 12F, 15B, 22F, and 33F), by Pfizer, are likely to

be licensed soon [6]. Additionally, a safe and immunogenic low-cost 10-valent vaccine for LMICs has been developed by the Serum Institute of India (SIIL-PCV), which also qualifies for AMC [206].

Serotype-agnostic vaccines, such as the killed whole-cell vaccine (WCV) or multi-component protein-purified vaccines (PPV), that in principal target all serotypes, are also being developed as a result of serotype replacement. These will likely not prevent NP colonisation and therefore will require different clinical trial endpoints to determine efficacy [207]. They are likely to be more affordable than PCVs as they do not require the conjugation process, something which is quite expensive. In addition to the previously mentioned pneumococcal surface proteins, the pillus of the bacteria have also been identified as potential targets for these vaccines [208]. Specifically, type 1 pillus, which was found in 25% of clinical isolates pre-vaccination and rebounded from reduced frequencies less than five years post-vaccination [209], increase host cell adherence and inflammatory responses [210]. However, its low prevalence in human populations and its bistable features (in which it is present in some clones and absent in others in an in vitro clonal population, i.e. phase variation) make its future as a vaccine candidate questionable [210]. Crucially, loci of some pneumococcal surface proteins, such as PspA and PspC, suggest high variability among the pneumococcal population, and therefore selection of protein antigens for serotype-agnostic vaccines requires careful consideration and continued genomic surveillance [24].

Despite the planned expansion of universally implemented PCVs, serotype replacement begs the question of whether PCV expansion is a sustainable goal, particularly given the consequences of the carriage-blocking vaccine mechanism, the geographic heterogeneity of the serotype landscape and the cost of developing such vaccines.

1.6 Thesis overview

In this thesis I will evaluate the impact of increasing the serotype coverage of PCVs in Europe and North America using epidemiological analysis methods based on mathematical models

and estimation of the key parameters that influence transmission and vaccine impact. The chapter structure and content is as follows:

Chapter 1 has been an introduction of the background of pneumococcus, including its epidemiology, pathology, detection methods, and control and interventions. Additionally, it has provided the motivation behind the thesis, as well as giving an overview of the thesis chapters.

Chapter 2 introduces dynamic transmission models and economic evaluations, and a systematic literature review of the published work and its limitations. This serves to identify the gaps in understanding in the current literature and to set the scene for the development of the mathematical model of infectious transmission and the impact of vaccination.

Chapter 3 provides an overview of serotype-specific IPD epidemiology in Europe, North America and Australia. This chapter presents an analysis of epidemiological trends and compares serotype-specific differences in post-vaccination IPD in a selection of countries with similar socioeconomic conditions, and describes the serotypes causing replacement disease that are potential additional targets for vaccine development.

Chapter 4 develops a novel methodology for estimating the case-to-carrier, or invasiveness, of pneumococcal serotypes in various settings before and after vaccination in children and adults. Since the mechanism of transmission is through carriage but carriage studies are time-consuming and expensive relative to national disease surveillance activity, invasiveness is essential in translating the number of carriers to the number of disease cases.

Chapter 5 describes the development and use of an individual-based stochastic model of transmission with a focus on estimating the impact of increasing serotype coverage in PCVs given the serotype-specific IPD epidemiology discussed in Chapter 3 and the invasiveness estimates evaluated in Chapter 4 in two example countries in Europe and North America: France and the United States.

Chapter 6 summarises the findings of each chapter and the contribution of the thesis. It also

discusses the limitations and extension of the results to countries other than those for which analyses and data were described. A key conclusion section is presented on the arguments for and against whether PCV valency should be increased universally. Finally, future directions for research are discussed.

Chapter 2

Systematic review of dynamic transmission models and economic evaluations of pneumococcal conjugate vaccines

This chapter presents a review of the published literature on transmission models of the impact of pneumococcal conjugate vaccines and economic evaluations comparing pneumococcal conjugate vaccines of increased valency. The aim is to explore the limitations of the published models and their assumptions and the implications of more robust methods for future vaccine formulation and policy.

2.1 Introduction

Mathematical models can provide insight into the transmission and control of infectious diseases, both to answer epidemiological questions and guide policy formulation for control [211]. The choice of model depends on the question being asked [212]. Dynamic (over time and space) and static (i.e. at equilibrium) analyses can be employed to examine specific problems. Models may be either deterministic or stochastic in form, and in the case of the latter tend to be based on individual-based simulations given the inherent difficulties in getting analytical insights from non-linear stochastic modelling studies.

The dynamics of transmission in infectious disease models are governed by the force of infection (FOI), defined as the per capita rate at which those who are susceptible become infected [211]. Because the FOI varies according to the number of people infected, it is not a fixed value and therefore dynamic transmission models (DTMs) are more appropriate for the study of infectious disease epidemiology and control. Static models, where the FOI is kept constant, are inappropriate in most cases except when considering equilibrium properties[212]. The FOI depends on the infectiousness of an infected individual. This parameter is influenced by many factors including the genetic backgrounds of host and pathogen. Infectiousness is typically positively correlated with viral or bacterial load, pathogen life cycle (e.g. is a free-living stage involved such as an airborne viral particle captured in exuded water droplets), and the host immune system, and host contact patterns (host behaviour) [213]. One of the most central concepts in infectious disease epidemiology is the basic reproduction number R_0, defined as the average number of secondary infections generated by one case in a large susceptible population [211]. If $R_0 > 1$, an infection can become an epidemic, whereas if $R_0 < 1$, infection will become extinct. Its value may vary from above unity (i.e. 1) to below unity according to season as is the case for the influenza A virus. The effective reproduction number R_t is defined as the average number of secondary infections from one case in a population at time t. The average number of secondary infections depends on the effective contact rate β, which is generally a function of biological infectiousness and host contact rates [213].

Changes in state variables can be modelled either deterministically or stochastically. In deterministic models, the outputs depend on input parameters and values, and the expected output is always the same. Deterministic models are therefore simpler than stochastic formulations, at the expense of epidemiological realism, but effective in estimating the coarser-grain dynamics of epidemics. Stochastic models have inherent randomness, such that the same input parameters and values will yield different results due to random number generation within the dynamics to decide what event occurs and when[213]. As such, they describe infectious disease systems more naturally [214], but are more computationally expensive and complex, and require a multitude of simulations to account for the variability between runs. Stochastic models can exhibit extinction behaviour in small populations regardless of the R_0, something which is not possible in deterministic models when the value of $R_0 > 1$ [215]. For most non-linear stochastic models, the expected mean is identical to the deterministic model prediction [216].

Models of infectious disease transmission and control can describe aggregate populations or if stochastic can track individual hosts over time. Individual-based models assign characteristics to each individual in the population employing defined behavioural rules and event rates. Aggregate population models split the population into compartments representing individuals that are susceptible (S), exposed (E), infected (I), or recovered (R). The compartments chosen depend on the properties of the infectious disease at hand, but the simplest model is SIR (Figure 2.1). If there is no immunity, then one might construct an SI or SIS model. For example, gonorrhoea might be modelled using an SIS structure because once the infection is cleared with treatment, an individual is susceptible again. Herpes simplex virus, on the other hand, can be modelled with an SI model since the host never clears the infection.
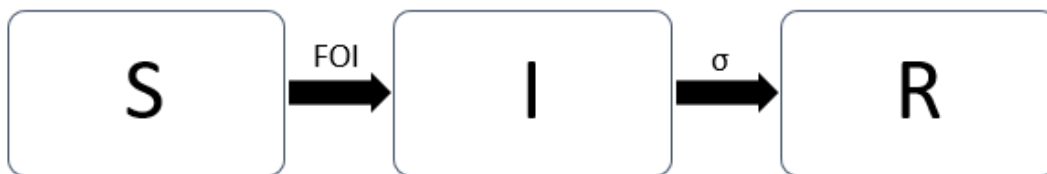
Figure 2.1: Simple Susceptible-Infected-Recovered model. FOI: force of infection; σ: recovery rate.

Moving from the susceptible compartment to the exposed or infected compartment depends on the FOI, which will depend on the number of infectious people, their contact rate with susceptibles, and the transmission probability on contact. From the infected state, individuals can move to the immune state (R) at the recovery rate, σ. In SIRS models, individuals can return to the S state at the rate of the loss of immunity. The expected number of each cohort can be estimated using ordinary or partial differential equations.

Models of infectious disease transmission are extremely useful for estimating the impact of interventions against infectious diseases, such as vaccination. In the case of compartmental transmission models, a vaccinated compartment can be added which individuals can move to at the rate of vaccination, and leave at the rate of vaccine waning. Vaccines can confer "leaky" or "all-or-nothing" protection, implying either a reduction in the risk of acquisition (or disease, depending on vaccine mechanism) in models, or a perfect blocking of acquisition.

Because transmission is dependent on the number of susceptibles, a high enough immune portion of the population can have indirect protective effects on the rest of the still susceptible population [217]. This phenomenon is referred to as herd immunity, which reduces the infection rate as the probability of contact between infected and susceptibles decreases. The proportion of the population that needs to be resistant to infection for this indirect effect to take place is referred to as the herd immunity threshold and is a critical aspect of evaluating vaccine impact [218]. To estimate the proportion of the population that would need to be vaccinated to block transmission, one can assume that in a homogeneously mixed population, $R_t = R_0 * x$, which must be below the unity level [218]. In this case, x is the fraction of the population that is susceptible. In a population where a proportion p are successfully vaccinated $x = 1 - p$. Rearranging:

$$R_t < 1 \tag{2.1}$$

$$R_0 * x < 1 \tag{2.2}$$

$$R_0 * (1 - p) < 1 \tag{2.3}$$

$$1 - p < \frac{1}{R_0} \tag{2.4}$$

$$p > 1 - \frac{1}{R_0}. \tag{2.5}$$

The critical percentage of the population that must be effectively immunised (taking into account vaccine coverage and vaccine efficacy) to reach the herd immunity threshold such that the outbreak would end is at least $1 - \frac{1}{R_0}$ [218]. In other words, the higher the R_0, the greater the fraction of the population that needs to be successfully vaccinated. In the context of PCVs whose primary target population is children, herd protection is established when enough of the child population is successfully immunised (i.e. accounting for the imperfect protection conferred by PCVs) against vaccine types (VT) such that these serotypes are eliminated from circulation, blocking further transmission and therefore reducing disease in the older populations not targeted by vaccines.

The development and analysis of mathematical models of multi-strain transmission and the impact of vaccination can help us understand the crucial relationship between vaccine-induced herd immunity and serotype replacement and help to include these indirect effects to examine post-vaccination transmission dynamics [219–221].

While several static transmission models have been developed to describe transmission and pneumococcal vaccine impact [222–226], these do not account for the indirect effects generated by herd immunity from those vaccinated and its concomitant effect on the likelihood of transmission to those still susceptible, and therefore are not able to capture the post-vaccination dynamics as effectively as DTMs. As DTMs aid us in estimating the indirect effects and disease reduction caused by vaccination, they can be extended to evaluate the economic benefit of choosing one vaccine over another.

Mathematical models can also be used to evaluate the value gained from implementing interventions against infectious diseases. Vaccines are generally cost-effective because they reduce disease, in turn reducing hospital visits, freeing up bed capacity and reducing the burden on caregivers [227]. However, choosing and adopting PCVs in infant immunisation

programmes requires a comparison of the benefits and costs associated with each vaccine. A standard approach is to conduct an economic evaluation (EE), which can be undertaken from various economic perspectives: the payer (i.e. healthcare or government) perspective (i.e. the public sector), the private perspective (i.e. the private sector), or the societal perspective. The societal perspective is considered the broadest, as it takes into account all foreseeable outcomes, including both direct and indirect costs. An example of an indirect cost would be the time lost from work due to illness, as this would result in loss of productivity. This would not be considered from a healthcare payer's perspective as this perspective only considers direct medical costs associated with illness.

There are several types of EEs, but cost-effectiveness analyses are those most frequently used, which express outcomes in natural units, such as the cost per infection averted. Cost-utility analyses are a type of cost-effectiveness analysis that quantify the effectiveness of different interventions using utility metrics, which relate to a person's level of well-being or health, and include quality adjusted life years (QALYs), disability adjusted life years (DALYs) or life years gained (LYG). Cost-utility analyses produce an incremental cost-effectiveness ratio (ICER) for two comparable interventions, which is the difference between their costs divided by the difference between their utility values. In essence, the ICER is used to estimate the additional cost for each additional health outcome, and is compared to a specific threshold to determine cost-effectiveness. This threshold of cost-effectiveness is somewhat contested but can be set by local guidelines (such as in the UK [228]), compared with other interventions, or in some cases based on a country's gross domestic product (GDP) per capita [229].

Health economic models are generally split into cohort or individual-level models, with the former evaluating the costs and utilities of proportions of a population following a health event, and the latter considering individuals with certain characteristics and being more computationally intensive. Cohort models are generally static, and the main types of cohort models are decision tree models (or decision-analytic models) and Markov models. Decision tree models occur instantaneously with different branches for each disease outcome, whereas Markov models have cycles over which the probability of a patient occupying a 'state' at

any given time is assessed [230, 231]. The approach chosen is rarely discussed in health economic papers [231]. In some cases, static models may be inappropriate for evaluating cost-effectiveness of interventions to prevent or control infectious diseases, as many of these interventions aim to reduce the FOI, thereby changing it [212].

I have conducted a systematic literature review to assess existing DTMs for their structure, data used, parameter estimation, general conclusions drawn, and finally the assumptions made on mixing, immunity, and serotype competition. This literature review aims to examine PCV DTMs and EEs for their potential contributions to informing public health immunisation policy, and to identify gaps in knowledge that can ultimately drive future directions of research.

2.2   Methods

Pubmed, Scopus, Ovid and ISI Web of Knowledge were searched up to October 5, 2021 using the following terms: PCV OR pneumococc* AND dynamic AND transmission AND model* AND (pneumococcal AND vaccine) AND (stochastic OR deterministic OR compartmental OR individual-based). The same databases, in addition to the Centre of Reviews and Dissemination (CRD), were searched up to October 9, 2021 using the following terms: "PCV OR pneumococc* AND (econ* OR cost benefit OR cost effectiveness OR cost utility OR economic impact) AND (pneumococcal AND vaccine)". Neither search was date restricted.

Published papers were included if:

- They presented a dynamic transmission model that investigated the epidemiological impact of any PCV in any setting OR they described any type of economic evaluation comparing PCVs in the European or North American setting (as defined by the World Health Organization)

-

- They were in English

- They were not a duplicate of another paper and were original models or evaluations

- A full article was available

A systematic literature review was conducted in accordance to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to investigate dynamic transmission models predicting PCV impact. Additionally, EEs were critically appraised for risk of bias using the Professional Society for Health Economics and Outcomes Research (ISPOR) checklist (Appendix A) [232, 233].

Data was extracted to Microsoft Excel to compare the published date, data source, demography/geography, interventions modelled, time horizon, model structure and type employed, immunity assumptions, disease states, mixing patterns, serotype classification or grouping, serotype competition, fitting process, sensitivity analysis, estimated parameters, and general findings. Due to the heterogeneity within each model aspect, it is difficult to standardize prediction values and compare the models quantitatively. Because of this, the publications were compared qualitatively.

2.3   Results

A total of 29 DTMs and 26 EEs were retrieved (Figure 2.2, Figure 2.3). These were evaluated for their characteristics (Table 2.1, Table 2.2) and summarized (Table 2.3, Table 2.4). Dynamic model heterogeneity highlighted the uncertainty around biological mechanisms. While all the EEs were deemed sufficiently relevant and credible (Appendix A.1), many had preventable weaknesses that could be avoided with increased transparency in their reporting of input parameters and conflicts of interest.
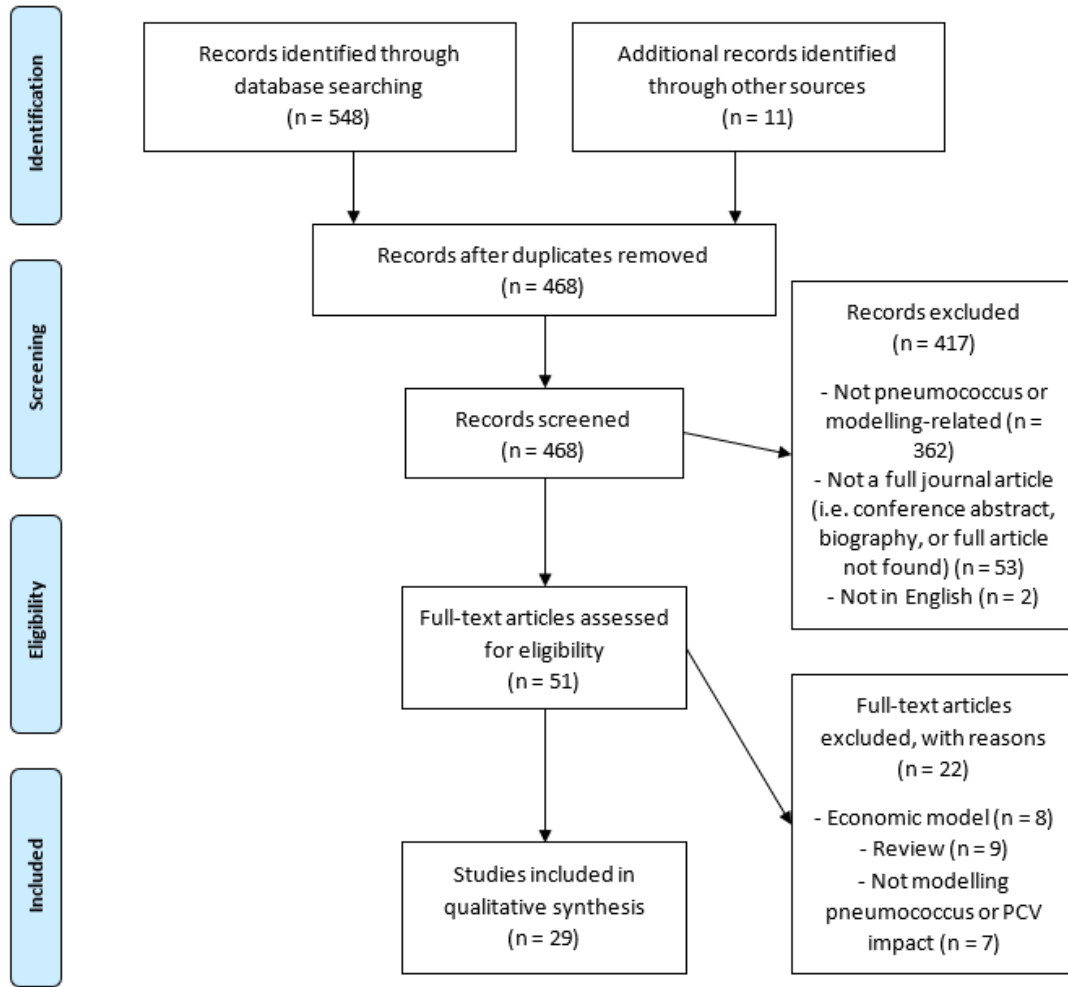
Figure 2.2: PRISMA flow diagram for dynamic transmission models review.
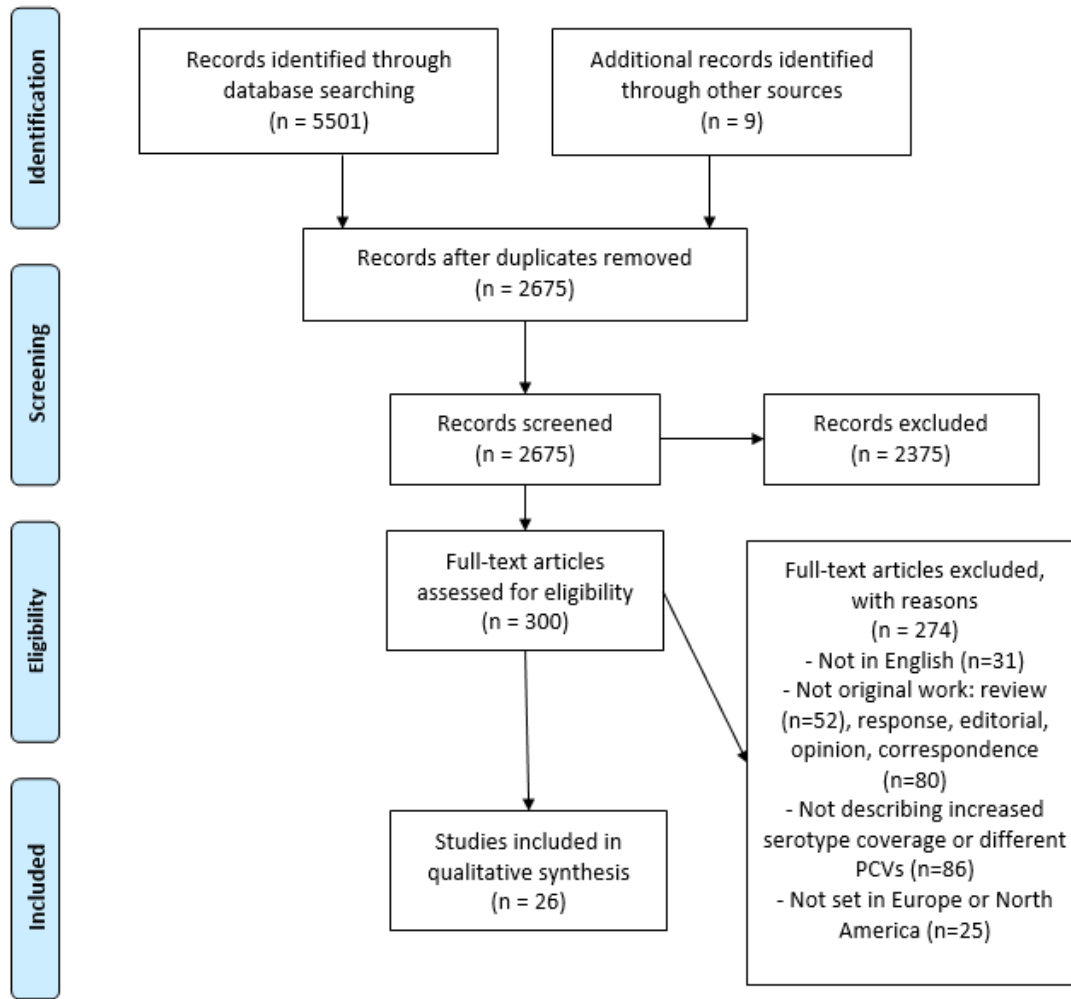
Figure 2.3: PRISMA flow diagram for economic evaluations review.

Table 2.1: Summary of dynamic transmission model characteristics. Analytical/mathematical models: focused on natural history of disease. Epidemiological models: focused on understanding the impact of vaccination. Policy-oriented: focused on vaccine policy.

Characteristic	Number of models (n = 29)	References
Type of model		
Compartmental		
Deterministic	23	[74, 84, 220, 221, 234–252]
Stochastic	1	[253]
Individual-based stochastic	4	[57, 254–256]
Hybrid	1	[257]
Scope of model		
Analytical/mathematical		
Without data	2	[84, 246]
Uses data without fitting	2	[242, 252]
Fitted to data	1	[57]
Epidemiological		
Uses data without fitting	4	[247, 249, 253, 254]
Fitted to data	12	[57, 220, 221, 234–237, 239, 243, 250, 255, 257]
Policy-oriented		
Uses data without fitting	1	[248]
Fitted to data	9	[74, 236, 238, 240, 241, 244, 245, 251, 256]
Parameter-fitting method		
Maximum likelihood estimation	2	[234, 255]
Nelder Mead method	1	[236]
Multinomial likelihood under Bayesian framework	4	[220, 239, 244]
Poisson deviance minimisation	5	[235, 237, 238, 243, 257]
Cross-sectional prevalence data	4	[220, 221, 238–240]
Minimisation of sum of differences of squares	3	[74, 250, 251]
Monte Carlo Markov Chain (MCMC)	4	[238, 241, 245, 256]
None	8	[84, 242, 246, 247, 249, 252–254]
Data used		
Uses pre-vaccination data	5	[57, 234, 241, 249, 256]
Uses post-vaccination data	2	[242, 252]
Uses pre- and post-vaccination data	20	[64, 74, 220, 221, 235–240, 243–245, 248, 250, 251, 253–255, 257]
Include economic evaluation	2	[240, 245]

Table 2.2: Summary of economic evaluation characteristics

Characteristic	Number of models (n = 26)	References
Type of model		
Static cohort		
Decision tree	15	[258–272]
Markov	7	[273–279]
Hybrid	2	[280, 281]
Dynamic	0	
Other	2	[282, 283]
Perspective		
Public sector only	10	[258, 259, 261, 263, 264, 269, 274, 275, 278, 279]
Private only	3	[267, 268, 283]
Societal (public and private)	13	[260, 262, 265, 266, 270–273, 276, 277, 280–282]
Analytical approaches		
Cost-effectiveness	24	[258–281]
Cost-benefit	2	[282, 283]
Indirect effects of vaccines		
Herd immunity	24	[258, 260–282]
Serotype replacement	18	[258, 262, 264, 265, 267–278, 280, 281]
PCV10 cross-protection of NVT	12	[259, 262, 265, 268, 271–278]
Protection against NTHi	16	[258, 259, 261–264, 266–268, 271, 273–278]
Includes transmission dynamics	0	
Fitted to data	0	

Table 2.3: Summary of included published dynamic transmission models

Authors and year published	Objective	Interventions modelled	Age / Age structure	Geography	Time horizon	Serotype classification[1]
SIS[2]						
Choi 2011	To predict the long-term impact of PCV7 on VT and NVT IPD incidence	PCV7 with different number of doses	Age-structured	High income country	50 years	VT, NVT
Choi 2012	To estimate effect on IPD cases of (i) replacing PCV7 with PCV13 (ii) discontinuing PCV vaccination altogether	Switch from PCV7 to PCV13, discontinuation of PCV, different number of doses	Age-structured	High income country	20 years post-vaccination	PCV7 VT, PCV13 VT, NVT
Choi 2019	To estimate the effect of changing the dose schedule from 2+1 to 1+1	PCV13; different dose schedule	Age-structured	High income country	15 years post-vaccination	PCV7 VT, PCV13 VT excl serotypes 1 and 3, NVT incl serotype 3
Cobey and Lipsitch 2012	To understand how relatively weak serotype-specific immune responses can support the coexistence of serotypes with substantial fitness differences	(PCV7)	Age-structured	High income country	150 years	By duration of carriage and in vivo competitive ability
De Cao 2014	To evaluate the potential effectiveness of PCV13 on IPD	PCV13	Age-structured	High income country	40 years post-vaccination	VT, NVT

Study	Objective	Intervention	Population	Setting	Time horizon	Serotypes
Flasche 2017	To estimate the dose efficiency of alternative catch-up campaigns in relation to PCV cohort introduction alone	PCV10 introduction scenarios with different catch-up strategies	Age-structured	Low income country	10 years	VT, NVT
Gjini 2016	To study how direct competition at co-colonization shapes multi-strain coexistence for neutral and non-neutral models	PCV7	Children 18-71 months	High income country	6 years post-vaccination	VT, NVT
Gjini 2017	To understand the epidemiological parameters responsible for differences in serotype replacement rates and to assess vaccine performance in field settings	PCV7	Pre-school/day-care children	High income countries	10 years	VT, NVT
Kitano and Aoki 2021	To understand the impact of the novel coronavirus (COVID-19) pandemic and the resulting disruption in childhood vaccinations on IPD incidence	PCV13	Age-structured	High income country	10 years	VT, NVT
Le Polain de Waroux 2018	To predict the impact of introducing PCV with and without a catch-up campaign	PCV13; catch-up campaign	Age-structured	Low income country	15 years post-vaccination	VT, NVT
Lipsitch 1997	To predict the factors affecting and the expected extent of serotype replacement in host population	No vaccine, serotype-specific vaccine, bivalent vaccine	N/A	N/A	N/A	VT, NVT

Study	Objective	Intervention	Model	Setting	Time horizon	Serotypes
Lipsitch 1999	To specify the conditions under which serotype replacement is likely, interpret results of PCV trials, design trials that would be better able to detect serotype replacement and suggest factors to consider in choosing serotype composition of vaccines	N/A	N/A	N/A	N/A	VT, NVT
Melegaro 2010	To predict the potential impact of PCV7 on the incidence of IPD	PCV7 and different schedules	Age-structured	High income country	50 years	VT, NVT
Mitchell 2015	To evaluate factors that may influence the emergence of a rare antibiotic-resistant types following vaccination targeting more common resistant types	PCV7	Children < 5 yrs	High income country	10 years	VT, NVT
Nurhonen 2013	To examine the net effectiveness of vaccination, the extent of replacement and herd immunity against VT carriage and IPD, and to compare effectiveness under different serotype compositions	PCV7, PCV10, PCV13	Age-structured	High income country	100 years	PCV7, PCV10 or PCV13 VT, and their case to carrier ratio
Ojal 2017	To predict the long-term impact of PCV10	PCV10 and catch up campaign	Age-structured	Low income country	10 years	PCV10 VT and 2 groups of NVT based on susceptibility to mutual competition

Study	Objective	Vaccine	Model structure	Setting	Time horizon	Serotype classification
Ojal 2019	To estimate whether continued vaccination is cost-effective after transition from Gavi funding	PCV13	Age-structured	Low income country	15 years post-vaccination	VT, strongly competitive NVT, weakly competitive NVT
Omori 2012	To clarify the effect of vaccine coverage and transmission dynamics on observed risk measures in the presence of cross-protective immunity between two strains	PCV7	N/A	N/A	N/A	VT, NVT
Scott 2014	To investigate how vaccine efficacy is best estimated from carriage data, at which time carriage should be assessed and to which factors timing is most sensitive	PCV7 and different schedules	Children < 5 yrs	Low income country	N/A	VT, NVT
Snedecor 2009	To quantify the direct and indirect benefits of vaccination	PCV7 and different number of doses	Age-structured	High income country	10 years post-vaccination	No classification - modelled PCV7 as a whole
Temime 2004	To investigate colonization in a vaccinated population and the effect of vaccine introduction on the distribution of resistance levels among carriers	PCV7, hypothetical PCV11	Age-structured	High income country	25 years	VT, NVT

Van Effelterre 2010	To evaluate the potential impact of antibiotic use and of vaccines with varying effects on serotype-specific IPD	PCV7, antibiotic treatment	Children <2 yrs	High income country	10 years	18 serotypes individually modelled, PCV7 VT, serotypes 6A and 19A, PCV7-related serotypes (except 6A and 19A)
Wasserman 2018	To quantify the increase in pneumococcal disease burden after the changed dose schedule from 2+1 to 1+1	PCV13; different dose schedule	Age-structured	High income country	10 years	PCV7 VT, serotype 3, serotype 19A, remaining PCV13 VT, NVT
Yang 2019	To understand the impact of background herd immunity on vaccine efficacy against carriage in individually randomized clinical trials of different dose schedules	PCV13; different dose schedule	Age-structured	High income country	105 years	Fitness of individual serotypes
SIS and **SI**[2]						
Gjini 2016	To estimate vaccine efficacy against acquisition and provide framework for parameter inference including direct competition	Theoretical	N/A	High income country	50 years post-vaccination	Pathogen 1 and Pathogen 2

SIR[2]

Study	Objective	Vaccine	Population	Setting	Time horizon	Stratification
Bottomley 2013	To predict the serotype replacement in carriage post-PCV introduction	PCV7	Children < 5 yrs	Low income country	10 years	By transmissibility and clearance rate (high, medium, low)

SIRS[2]

Study	Objective	Vaccine	Population	Setting	Time horizon	Stratification
Flasche 2013	To assess the impact of vaccination with a bivalent and a universal vaccine, and both in succession	Bivalent vaccine and universal vaccine (immunity against all serotypes)	Age-structured	High income country	30 years	By intrinsic duration of carriage
Zhang 2004	To study how direct and indirect competition and vaccination influence the coexistence of two serotypes in a population	N/A	N/A	N/A	N/A	VT, NVT

SEI[2]

Study	Objective	Vaccine	Population	Setting	Time horizon	Stratification
Sutton 2008	To assess the impact of a program providing pneumococcal vaccines to large risk groups	PCV7, PPV23	Age-structured	High income country	5 years	N/A

[1] **VT**: Vaccine-type; **NVT**: Non-vaccine type.
[2] **SIR**: Susceptible-Infected-Recovered; **SIS**: Susceptible-Infected-Susceptible; **SIRS**: Susceptible-Infected-Recovered-Susceptible; **SI**: Susceptible-Infected; **SEI**: Susceptible-Exposed-Infected.

61

Table 2.4: Summary of included published economic evaluations categorized by outcomes assessed

Authors and year published	Interventions	CE results	Cohort / Population-based model type	Country	Diseases[1]	Potential conflict of interest
QALY, LYG, DALY						
Beutels 2011	PCV7 vs PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	Belgium	AOM, PNE[2] MEN, BAC, other IPD	Multiple authors had grants from both GSK and Pfizer
QALY, LYG						
Bakir 2012	PCV7 vs PCV10 vs PCV13	PCV10 more cost-effective than PCV7 and PCV13	Decision-analytic	Turkey	MEN, BAC, PNE[2] (CAP), AOM	Funded by GSK; one author is an employee of GSK
Blank 2012	PCV7 vs PCV13	PCV13 more cost-effective than PCV7	Decision-analytic	Switzerland	MEN, BAC, PNE[2] AOM	Funded by Pfizer
Castiglia 2017	No vaccination vs PCV7 vs PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	Italy	MEN, BAC, PNE[2] AOM	Funded by GSK
De Wals 2009	PCV7 vs PCV10	PCV10 more cost-effective than PCV7	Decision-analytic	Canada	IPD[2] AOM, PNE[2]	Funded by GSK
Earnshaw 2012	PCV10 vs PCV13	PCV13 more cost-effective than PCV10	Decision-analytic	Canada	IPD, PNE[2] AOM	Funded by Pfizer

Kuhlmann 2017	PCV7 vs PCV10 vs PCV13	PCV13 more cost-effective than PCV10	Decision-analytic Markov hybrid	Germany	IPD (MEN + BAC), PNE^2 AOM	No conflict of interest reported
Rozenbaum 2010	No vaccination vs PCV7 vs PCV10 vs PCV13	PCV13 most cost-effective and PCV7 least cost-effective vs no vaccination	Decision-analytic	Netherlands	MEN, BAC, PNE (invasive and non-invasive1), AOM	Authors have grants from Wyeth/Pfizer/GSK/Novartis/Baxter
Rubin 2010	PCV7 vs PCV13	PCV13 cost-saving vs PCV7	Decision-analytic Markov hybrid	USA	IPD (MEN + BAC), PNE^2 AOM	Funded by Wyeth/Pfizer
Strutton 2012	PCV7 vs PCV10 vs PCV13	PCV13 more cost-effective than PCV7 and PCV10	Decision-analytic	Germany, Greece and Netherlands	IPD, PNE^2 AOM	Funded by Wyeth/Pfizer
Talbird 2010	PCV7 vs PCV10	PCV10 more cost-effective than PCV7	Decision-analytic	Canada, Mexico, Germany, Norway	MEN, BAC, PNE^2 AOM	Funded by GSK
QALY						
Ansaldi 2020	PCV10 vs PCV13	PCV13 more cost-effective than PCV13	Decision-analytic	Italy	IPD (MEN + BAC), PNE^2 AOM	Funded by Pfizer, three authors employees of Pfizer
By 2012	PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	Sweden	IPD (MEN+BAC), PNE^2 AOM	Funded by GSK, two authors employees of GSK

Study	Comparison	Result	Model	Country	Outcomes	Conflict of interest
Chuck 2010	PCV7 vs PCV10 vs PCV13	PCV13 more cost-effective than PCV10, unless effect against NTHi is included	Decision-analytic	Canada	MEN, BAC[2] PNE[2] (invasive and non-invasive), AOM	No conflict of interest reported
Delgeize 2016	PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	UK	IPD (MEN+BAC), PNE[2] AOM	Funded by GSK
Klok 2013	PCV10 vs PCV13	PCV13 more cost-effective than PCV10	Decision-analytic	Denmark, Sweden	AOM, PNE[2] IPD	All authors employees of Pfizer
Knerer 2012	PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	Canada, UK	MEN, BAC[2] PNE[2] AOM	All authors employees of GSK
Pugh 2020	PCV10 vs PCV13	PCV13 more cost-effective than PCV10	Decision-analytic	Finland, Netherlands (and Colombia)	IPD (MEN + BAC) PNE[2] AOM	All authors employees of Pfizer
Robberstad 2011	PCV7 vs PCV10 vs PCV13	PCV10 more cost-effective than PCV13	Markov	Norway	MEN, BAC, PNE[2] AOM	One author employee of GSK, one with grant from GSK
Smith 2021	PCV13 vs PCV15 vs PCV20 vs hypothetical PCV20	Current PCVs in pipeline not cost-effective in older adults	Markov	USA	IPD, PNE	Reported conflicts of interest unrelated to topic
Vemer 2014	PCV10 vs PCV13	PCV13 more cost-effective than PCV10 if indirect effects considered	Decision-analytic	Netherlands	AOM, PNE[2] IPD	Funded by GSK

64

Wilson 2018	PCV10 vs PCV13	PCV13 more cost-effective than PCV10	Decision-analytic	Canada	IPD (MEN + BAC), PNE[2] AOM	Funded by Pfizer
DALY						
Nakamura 2011	No vaccination vs PCV7 vs PCV10 vs PCV13	All PCV would be highly cost-effective, PCV13 cost-effective for more middle-income countries	Decision-analytic	Middle-income countries	PNE[2] MEN, NPNM	One author's husband owns Pfizer stock, another had grant from Pfizer
Vucina 2015	No vaccination vs PCV10 vs PCV13	Neither PCV considered cost effective	Decision-analytic	Croatia	AOM, PNE, MEN, NPNM[2]	No conflict of interest reported
Cases averted						
Luca 2018	PCV7 vs PCV10 vs PCV13	PCV13 saves money and prevents further pneumonia hospitalisations	Differences-in-differences approach	Canada	PNE, Other (MEN, NPNM, SEP, EMP)	No conflict of interest reported
Waye 2015	PCV7 vs PCV13 (+ PPV23)	PCV13 saves money and prevents further IPD cases	Economic costing model	Canada	PNE, BAC[2] MEN	No conflict of interest reported

[1] AOM: acute otitis media. BAC: bacteraemia. CAP: community-acquired pneumonia. EMP: empyema. LY: life years. MEN: meningitis. NPNM: Pneumococcal non-pneumonia non-meningitis. NTHi: non-typeable Haemophilus influenzae. PNE: pneumonia. QALY: quality-adjusted life years. SEP: septicaemia.
[2] Both inpatient (hospitalized) and outpatient

2.3.1 Dynamic Transmission Models

Data from high-income countries is widely used to parameterise these models because introduction of PCV into low-income countries has lagged behind more developed countries [284], even though the disease burden is higher in low-income countries [102]. Additionally, many of the published papers evaluated PCV7 rather than PCV10 or PCV13 because the latter did not receive marketing approval until later [285, 286]. Although most of the models used both pre- and post-vaccination data (Table 2.1) for model parametrisation and fitting, models using only pre- [57, 234, 241, 249, 256] or only post-vaccination data [242, 252] were limited in their ability to accurately reproduce a setting's transmission dynamics.

The inconsistency in model structure highlights the complexity of naturally-acquired immunity against pneumococcus, and the scarce evidence of its impact or how it persists post-recovery from infection. Many of the papers used a deterministic SIS model structure, while others employed deterministic SIR, SEI, SI, and SIRS models. Cross-immunity was not accounted for in any models but three [57, 234, 246, 254, 256], which considered non-specific immunity to be transient. Vaccine-induced immunity, on the other hand, was assumed to be either leaky or "all or nothing". The actual biological action of either is largely unknown at present. Additionally, most models were deterministic, with only six (one of which was a hybrid) being stochastic models. Some models were age-structured, splitting the population into various age groups. Without age-structure, models ignore the impact that age groups could have on others. Children have a higher carriage prevalence than adults, particularly those living with siblings or attending day care centres [6]. This known biological feature cannot be reflected in models without age-structure.

Country-level differences may have impacted the results of the models. Population size was predicted to have minimal to no impact in some cases [57, 244, 254], but decreased the vaccine efficacy in others [221]. On occasion, models were insensitive to the population structure (or mixing) [57] but very sensitive in others [237, 243]. These contradictions reveal an uncertainty in the significance of these factors on vaccination impact. Gjini's model showed that increasing heterogeneity within a population reduced the impact of vaccination [220]. It also showed that

vaccine efficacy and prevalence changed depending on the country on which it was employed because of differing serotype distributions [220]. Two models found that increasing the catch-up campaign reduced the IPD incidence much faster [241, 243]. Two models evaluated a reduced dose schedule in England and Wales (2+1 to 1+1), finding conflicting results. One, conducted by Public Health England, found that the vaccine impact would still be largely preserved [236], whereas the other, conducted by the vaccine manufacturer, found that the disease burden would increase substantially [251]. The latter did not fit to carriage studies, but rather to outdated case-to-carrier ratios (CCRs). Countries with variation in vaccination schedules and coverage vary in vaccine intensity, and thus herd immunity. This is further demonstrated in one of the models [256] which found that vaccine efficacy in dose schedule clinical trials changes depending on context and timing, and therefore are only comparable in vaccine-naive populations.

Even though the outcome of transmission models is usually disease cases, pneumococcal transmission is through carriage, and it is also a precursor to disease [70]. Despite this, only two of the models included both colonized and IPD states [74, 287], and one of these [287] described a system in which carriage was not a precursor to disease. Most other studies used CCRs to estimate the disease cases from the number of carriers. Clinical epidemiological studies suggest that up to six different serotypes can be carried concurrently [169], although co-colonization becomes increasingly rare with more serotypes being carried due to the competitive interaction between serotypes [288]. As such, all the models except three [236, 253, 257] assumed that more than two serotypes could not be carried simultaneously. Detection of carriage is biased towards single carriage reporting, but only one model accounted for this, resulting in a decreased competition parameter and an increased reproductive number [220, 239]. Multiple carriage is important to consider for pneumococcal colonisation and competition dynamics, and thereby serotype replacement. Vaccination prevents carriage of VT, but it would be interesting to gather more data on whether multi-serotype coexistence affects the impact of PCV or vice versa.

The dominating assumption in the retrieved DTMs is that individual serotype characteristics

are vague and generalised. Other than a few models that grouped serotypes by biological and epidemiological characteristics, most studies classified serotypes into two supergroups as either VT or NVT. When the purpose of the model was to assess PCV vaccination or understand underlying epidemiological mechanisms such as colonisation, immunity, competition or replacement, this classification may not be appropriate. Having two supergroups oversimplifies the system and loses details on carriage and disease patterns, as well as serotype interactions and consequently serotype replacement. In a similar vein, serotype competition was assumed to reduce the acquisition rate of a new serotype when the human host is already colonised, but with only two supergroups, this also loses the granularity of serotype-specific replacement. Neutral models especially [57, 74, 220, 234, 235, 240, 243, 253, 255, 289–291], defined as having equivalent transmission and clearance rates across all serotypes, fail to provide details of serotype-specific transmission intensity, clearance rates or duration of carriage, offered by their serotype-specific counterparts [234, 249, 250, 254, 255]. One reason for this is the limited data on the natural history of infection. This is reflected in the range of parameter values, as well as contradictions in sensitivity analyses on factors such as duration of vaccine protection. Numerous studies have established that a host's susceptibility to disease is mainly associated with acquisition of a serotype rather than a prolonged duration of carriage [70], so a serotype's natural history of infection may dictate its fitness in terms of competition within an individual and therefore its prevalence in the population. Two models classified serotypes in this fashion, with observed prevalence of serotypes dictating serotype-specific fitness [57, 256]. Understanding between-serotype competition within an individual would enable us to model serotype replacement and herd immunity more effectively, and potentially allow us to predict which serotypes are more cause for concern and should be targeted by PCV.

2.3.2 Economic Evaluations

The main differences between the included economic models were their assumptions on four factors, namely: herd immunity, serotype replacement, PCV10 impact on AOM caused

by non-typeable *Haemophilus influenzae* (NTHi), and PCV10 cross-protection against NVTs
(Table 2.5).

Of the included economic papers, most described a European setting, and the remainder
a North American setting. They typically employed epidemiological data on surveillance,
serotype profiles, vaccine efficacy, NIPD distribution, and indirect effects. Indirect effects
were not calculated based on transmission dynamic models that captured herd immunity, and
as such all models were static. Early pneumococcal economic evaluations were static as it was
unknown whether vaccines prevented transmission or simply blocked disease [212]. Because
of their relatively recent licensure, much of the data on PCV10 and PCV13 used by the models
was extrapolated from PCV7 data. PCV10 and PCV13 clinical trial data have demonstrated
immunogenicity, safety and tolerability rather than clinical effectiveness. Currently, the
latter is inferred from earlier vaccine formulations. This demonstrates a limitation of the
studies. Despite this, most papers compared PCV10 and PCV13. Only one, more recent study,
compared PCV13 to PCV15 (PCV13 + 22F and 33F), PCV20 (PCV15 + 8, 10A, 11A, 12F, and
15B) and a hypothetical PCV20 based on serotypes causing disease in older adults [279].

Table 2.5: Summary of assumptions on indirect effects, NVT cross-protection, and NTHi for each economic evaluation

Authors and year published	Herd immunity included	Serotype replacement included	Cross-protection of NVT	Protection against NTHi included
Ansaldi 2020	Implicitly included in indirect effects for IPD	Yes: for IPD only	No	Yes: for both PCV10 (21.5%) [186] and PCV13 (24.5%)
Bakir 2012	No	No	Yes: against 6A	Yes: POET study (35.6%) [187]
Beutels 2011	Implicitly included in net indirect effects: reduced number of disease cases. Considered with and without in scenario analysis	Implicitly included in net indirect effects: considered with and without in scenario analysis (0-99%)	Yes: with and without protection against 19A	Yes: POET study (34%) [187]
Blank 2012	Yes: reduction in mortality rate of unvaccinated people	No	PCV10 not considered	PCV10 not considered
By 2012	Implicitly included in net indirect effects for IPD only	Implicitly included in net indirect effects for IPD only	Yes: with and without protection against 6A (76%) and 19A (26%) IPD	Yes: POET study (35.6%) [187]
Castiglia 2017	Implicitly included in overall vaccine effectiveness	Implicitly included in overall vaccine effectiveness	Yes: against 6A and 19A	Yes: COMPAS study (21.5%) [186]
Chuck 2010	Yes: with and without	No	No	Yes: with (additional 5% reduction) and without

Authors and year published	Herd immunity included	Serotype replacement included	Cross-protection of NVT	Protection against NTHi included
De Wals 2009	Yes: for IPD only (not PNE and AOM)	Yes: for IPD only (not PNE and AOM)	Yes: against 6A and 19A	Yes: POET study (35.6%) [187]
Delgeize 2016	Implicitly included in indirect effect for IPD only: 30% reduction in disease incidence	Implicitly included in indirect effect for IPD only: 30% reduction in disease incidence	Yes: against 6A and 19A	Yes: COMPAS study (21.5%) [186]
Earnshaw 2012	None for PCV10; PCV13 assumed to have same indirect effects (herd protection) as PCV7	Not explicitly examined	No	Yes: 4%
Klok 2013	Implicitly included in net indirect effects	Implicitly included in net indirect effects	No	Yes: 4%
Knerer 2012	Implicitly included in net indirect effects for IPD only; 15.4% (< 5 yrs) and 29% (> 5 yrs)	Implicitly included in net indirect effects for IPD only; 15.4% (< 5 yrs) and 29% (> 5 yrs)	Yes: against 6A (AOM, IPD) and 19A (IPD)	Yes: POET study based on PCV11 predecessor (35.6%) [187]
Kuhlmann 2017	Implicitly included in net indirect effects for IPD and NIPD	Direct vaccine effect was reduced proportionally with the decrease in direct effects against IPD due to replacement disease	No	No
Luca 2018	Implicitly included in indirect effects	Not explicitly examined	No	No
Nakamura 2011	Yes: 50% of US herd protection	Yes: - assumed same as in USA	Yes: against 6A	No

Authors and year published	Herd immunity included	Serotype replacement included	Cross-protection of NVT	Protection against NTHi included
Pugh 2020	Implicitly included in indirect effects	Implicitly included in indirect effects	No	Yes: for both PCV10 (21.5%) [186] and PCV13 (24.5%)
Robberstad 2011	Implicitly included in net indirect effects for IPD only	Implicitly included in net indirect effects for IPD only	Yes: with and without protection against 6A (76%) and 19A (26%)	Yes: with (35.6%) and without
Rozenbaum 2010	Yes: for IPD only (unvaccinated children, and vaccinated but not yet protected), 10%	Assumed included in vaccine efficacy estimates	No	No
Rubin 2010	Implicitly included in net indirect effects	Implicitly included in net indirect effects	PCV10 not considered	No
Smith 2021	Implicitly included in indirect effects	No	PCV10 not considered	No
Strutton 2012	Implicitly included in net indirect effects: considered PCV10 and PCV13 with and without in scenario analysis	Implicitly included in net indirect effects: considered PCV10 and PCV13 with and without in scenario analysis	No	Yes: 4%
Talbird 2010	Implicitly included in net indirect effects: considered with and without in scenario analysis	Implicitly included in net indirect effects: considered with and without in scenario analysis	Yes: against 6A and 19A	Yes: POET study (35.6%) [187]

72

Authors and year published	Herd immunity included	Serotype replacement included	Cross-protection of NVT	Protection against NTHi included
Vemer 2014	Implicitly included in indirect effect: considered with and without in scenario analysis varying percentage of UK indirect effects	Implicitly included in indirect effect: considered with and without in scenario analysis varying percentage of UK indirect effects	Yes: against 19A	Yes: with POET (35.6%) [187] and COMPAS (21.5%) [186] studies and without
Vucina 2015	Considered in scenario analysis	Yes	No	No
Waye 2015	Not explicitly examined	No	PCV10 not considered	PCV10 not considered
Wilson 2018	Implicitly included in indirect effects	Yes	Implicitly included in indirect effects	No

Country guidelines for health economic perspectives [292] were not necessarily followed [269, 275], particularly by studies comparing the same perspective for different countries [267, 268, 276]. However, the studies that included multiple health economic perspectives [262, 270, 276, 280] generally had consistent results on which vaccine was more cost-effective.

Interestingly, the included EEs were nearly all industry studies, which at times fundamentally biased the parameters used (Table 2.4, Appendix A.1), and thus the interpretation of the results. A majority reported conflict of interest with either GSK or Pfizer, the two competing PCV manufacturers who funded or provided grants for these studies or employed one or more of the authors. The vaccine that was manufactured by the funding or supporting company was always preferred or considered more cost-effective. Of the six papers that did not report a conflict of interest, three reported that PCV13 was more cost-effective [280, 282, 283], and one reported that PCV10 was considered more cost-effective if the impact on NTHi was included [261]. One reported that neither were cost-effective, although it is worth noting that this study did not include indirect effects in their analysis [270]. The last one compared PCV13 to higher valent vaccines that are not currently licensed (PCV15 and PCV20) in older adults and found that the PCVs based on children's IPD would not be cost-effective in adults [279]. Additional published work from national public health agencies, or detailed methodology including the completion of a reporting guidance checklist, such as the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist, could help eliminate the bias and make the results more transparent and robust.

One of the challenges in comparing these studies is that they all described different settings, with different currencies and price years (Appendix A.2). Costs of vaccines as well as hospitalization for IPD and AOM differed between countries, but also within countries. For example, EEs for Germany [267, 268, 280] reported different costs of hospitalization for the same price year, resulting in different ICERs. Regardless, PCV13 was reported as cost-effective compared to its lower valency counterparts in these studies.

Herd immunity and serotype replacement were often grouped together as a 'net indirect

effect' that was fixed to a specified range of values. When this was the case, it was applied to disease as a reduction in disease incidence by age group. Herd immunity, considered separately, was applied to non-vaccinated populations as a result of the decreased disease circulation or mortality rate. This is considered acceptable when herd immunity is based on high quality surveillance data, although there is no gold standard or formal guideline for incorporating herd immunity into static EEs [293]. It was sometimes also considered inherent to the data analysed [258, 266, 272, 279, 282]. In all cases, indirect effects were not calculated based on a transmission dynamic model framework. Serotype replacement alone was excluded in three of the studies that considered the two effects separately [260, 261, 263]. In half of the sensitivity analyses, indirect effects were a key driver of, or had a moderate impact on, cost-effectiveness [260, 263–267, 270, 271, 278–281]. Indirect effects vary over time as more people infected acquire immunity. The range of fixed values used by these economic models for indirect effects were rarely calculated based on vaccine coverage or the duration of the vaccine programme. One study assumed the indirect effects were implicit based on historical trends used to forecast IPD, though historical trends may not be useful in predicting serotype replacement. While fixed values may be acceptable when based on quality surveillance data for short time horizons, these might not be extendable to determine the long-term impact [293]. Consequently, the impact of indirect effects is difficult to deduce as the models were static.

NTHi reportedly accounts for approximately a third of AOM cases [185]. When PCV10 was evaluated, its impact on NTHi-related AOM was not always included. This impact is controversial [187, 294, 295]. The magnitude of this protection in the studies was assumed to range from 4% to 36% from either the Pneumococcal Otitis Efficacy Trial (POET) [187] or the Clinical Otitis Media and Pneumonia Study (COMPAS) [186]. Generally, GSK-supported papers added a higher protection against it (35.6%), from the POET study's PCV10 efficacy against clinical AOM episodes caused by *H. influenzae* rather than the reported value of 35.3% for NTHi. On the other hand, Pfizer-supported studies ignored it or added minimal protection (4%), this value referring to the POET study's PCV10 efficacy against the first episode of AOM

caused by NTHi, multiplied by the fraction of cases in the control AOM group that were NTHi-caused. PCV13 may prevent NTHi-related AOM caused by complications from infection by VTs [296], and as such, one Pfizer-supported study included PCV13 protection against NTHi (24.5%) [258], though no reference for this specific value could be found. Regardless, AOM-related outcomes were a key driver of sensitivity in many models [259, 260, 269, 273–278], and as such this could be an important introduction of bias in the studies.

Cost-effectiveness was also driven by PCV10's cross-protection against NVTs. Most Pfizer-supported studies did not include any explicit PCV10 cross-protection against NVTs [258, 263, 264, 266, 267, 269, 270, 280], while all the GSK-supported studies did [259, 262, 268, 271, 273–277]. There are several epidemiological studies that support PCV10 cross-protection against serotypes 6A and 19A [186, 187, 297–301], although this is somewhat controversial depending on the setting. Despite this, sensitivity analyses showed that even if excluded, the conclusions were the same [271, 273, 277, 278]. Even though serotype 19A has been one of the most prevalent serotypes around the world [194, 198, 302–305], PCV10's cross-protective efficacy against it is only 26% [297], a possible explanation for why its inclusion did not alter model conclusions.

2.4 Discussion

This review gathered PCV DTMs and EEs and highlighted some of the gaps in our knowledge in the epidemiological aspects of vaccination. DTMs generally found PCVs to be effective in reducing the burden of IPD and EEs found PCVs to be cost-effective, although both sets of papers made different assumptions which suggests biological uncertainties. In both cases, the results relied primarily on herd immunity and serotype replacement which are difficult to parameterise given their dynamic nature with time and heterogeneous populations.

Indirect effects are crucial to vaccine evaluation and cannot be captured as accurately by static models. DTMs show that PCV7 is an effective way to reduce the disease burden although they still cannot quantify serotype replacement yet, particularly as a majority still group serotypes

into VT and NVT supergroups. One way to circumvent this and be computationally feasible while still capturing the indirect effects might be to group serotypes according to serotype characteristics, such as acquisition and clearance rates, as one of the models did [234]. How and whether one chooses to group serotypes will depend on the goal of the study. A model investigating the mechanism of competition might only need two serotypes, whereas a model investigating the effects of vaccination or multi-serotype coexistence and the mechanism of serotype replacement may require more. Additional biological and epidemiological data, such as on immunity (natural, vaccine-induced, and cross-serotype) and duration of protection would allow for parameter estimation on herd immunity and serotype replacement, especially for PCV10 and PCV13. Additionally, interaction between individuals (age specific rates of infection and mixing patterns between age groups), something that is essential when modelling infectious disease transmission and vaccine impact, were often excluded, as was an interaction between serotypes (competition). Several mechanisms of competition have been described [56, 306–313], but the precise mechanisms and intensity of inhibition by direct competition remain to be explored. One postulated mechanism, for example, states that serotypes requiring less in a nutrient-deficient environment (i.e. are less metabolically demanding) are those that are more prevalent and therefore those considered more "fit" [62, 288, 314]. Competition parameters cannot be directly measured but are currently extrapolated from surveillance data. Large sample sizes are required in surveillance data that models rely on to implement an age-dependent multiple carriage prevalence, which is rarely the case in practice [257]. This would allow DTMs to better estimate the disease burden after vaccination and to predict which serotypes are more crucial to include in the vaccine. This data would also add robustness in the EEs such that bias from conflicts of interest would not prevail. Static EE models simplify heterogeneity and do not allow an estimation of the changes in herd immunity or serotype replacement over time, particularly as vaccine coverage increases. Static EEs may be warranted if the time horizon is shorter, however serotype replacement has a long-term impact. They overestimate effectiveness if, with increasing age at infection as a consequence of increasing vaccine coverage, the risk of developing complications rises as is recorded in a variety of epidemiological studies [315]. Static economic models are

also unsuitable as they generally can underestimate the impact of interventions such as vaccines, which reduce levels of infection transmission in a population [316]. These economic models are far too simplistic at present in the absence of a transmission dynamic framework. Interaction effects are not explicitly considered in static models, but they are in DTMs, even if much of the data required to precisely parameterise these effects is scarce. Two of the models included in the DTM review did include an EE as well [240, 245]. However, they were not included in the EE review because they did not meet the criteria of comparing PCVs, nor were their settings European or North American.

At the core of pneumococcal transmission dynamics is the presence of multiple serotype carriage. Currently only bi-carriage has been modelled. The consequence of a higher multiple serotype carriage on transmission dynamics is not yet understood and adds a degree of complexity which may be unnecessary depending on the model's objective, with the need to estimate many more parameters especially in connection with cross serotype immunity. Regardless, used in conjunction with an invasiveness estimate (case-to-carrier ratio) to extract the number of disease cases. Should the data become more abundant, one could also extract the number of specific disease outcomes (meningitis, sepsis, etc) per serotype, which could link the DTMs to the EEs.

The reviewed transmission models never accounted for NIPD, whereas it was often included in EEs and assumed to have a similar serotype profile to IPD. This may be because the PCV impact on NIPD is not measured or required for the licensure [70], and because serotype-specific data on AOM and pneumonia is sparse and therefore it is difficult to predict long-term effects of vaccination or indirect effects. Nevertheless, AOM accounts for most pneumococcal disease costs because of its frequency, is important in regards to its contribution to antimicrobial resistance [317]. As a key driver in the published sensitivity analyses, its parameters merit exploration and quantification.

The population-level differences between countries signify that basing data from one country makes predictions specific to this region. It is unknown at present whether these predictions

apply to other populations [250]. Higher carriage and IPD prevalence in, for example, developing countries and native population settings may be due to less effective immune responses arising from malnutrition, genetic differences, or other epidemiological risk factors such as HIV infection [104, 105] as well as from increased number of contacts [254]. These factors affect the case-to-carrier ratios, which are age- and serotype-dependent, and could be used to directly measure IPD prevalence. Furthermore, increasing the transmission intensity might increase the occurrence of multiple carriage in a host population, which would increase the indirect effects of competition [221]. In the EEs, most models assumed price parity between vaccines even though this is highly unlikely in practice. Cost (vaccine dose price, medical costs), a main driver in EEs, and different tendering outcomes, would change the conclusions derived from economic models specific to one country. The extent of the benefits of PCVs on populations of different sizes, age structures, or where differing pricing policies prevail will depend on their respective epidemiological and economic characteristics. As such, predictions may not currently be easily comparable.

While much past research has focused on DTMs and EEs separately, there has been little focus on developing EEs based on DTMs and only two DTMs including an EE in this field [240, 245]. DTMs provide a template for economic analyses of the relative benefit of adding more serotypes to the conjugate vaccine. This cannot happen if they only model two supergroups. Assessing vaccine impact in terms of both carriage and disease provides a greater understanding of the transmission dynamics that can ultimately improve our ability to make informed policy decisions [70]. However, doing this effectively is challenging at present given our poor understanding of the relationship between pneumococcal carriage and disease, including multi-serotype coexistence, and PCV's role in it. Until we have a better understanding of these factors, EEs will not be able to reflect the true value of implementing one PCV over another.

2.5 Aim in this thesis

I have identified a series of areas where current understanding of the impact of pneumococcal conjugate vaccines is very poor. As such, this presents many challenges to the development of pneumococcal transmission models (and associated parameter estimation) and EEs of the vaccines' impact. Specifically, many questions remain about how a multi-serotype dynamic affects the natural history of infection and the occurrence of associated disease. The next chapter will focus on understanding the trends of disease serotype replacement in countries having implemented PCVs, to identify serotypes that replace previously dominant types in disease post-vaccination in these countries and whether such settings would benefit from a vaccine targeting the replacing serotypes.

Chapter 3

Serotype replacement trends in high-income countries and significance to expanding vaccine valency

This chapter gives an overview of serotype replacement trends in invasive pneumococcal disease in Australia, the United States and select European countries before and after vaccination with PCV. The aim was to understand how the serotypes responsible for disease after vaccine introduction in high income countries with similar demographics has changed and, as a result, the case for developing vaccines with expanded valency.

3.1 Introduction

As expanded PCVs are likely to be licensed soon [6], it is in the interest of the public health community to understand which serotypes are causing disease in countries where vaccines have been introduced and whether there is commonality in replacement patterns across countries. After PCV7 was implemented, serotype 19A became highly prevalent in carriage and disease worldwide [167, 191, 192, 195, 318–321], leading to its inclusion in the higher valent PCV13. An important question therefore is whether the serotype landscape post-PCV10/13 is again amenable to the targeting of a select few additional serotypes with an expanded valency vaccine. While such a vaccine would be implemented globally, understanding the trends in high-income countries that have already implemented higher valent PCVs and that have similar socioeconomic and demographic characteristics can give a preliminary indication of whether this undertaking would be beneficial to population health.

This chapter has four aims. First, to assess the patterns of relative serotype abundance in IPD cases before and after the introduction of PCVs in North America, Australia and selected countries in Europe. Second, to examine differences by country in these patterns in relation to the vaccines introduced in national immunisation programmes. Third, to assess the expansion of replacement serotypes in comparison with the serotypes targeted by the vaccine. The final aim is to identify potential candidates for serotypes that should be included in conjugate vaccines given increasing evidence on the invasiveness of the replacement serotypes.

3.2 Results

3.2.1 Dataset characteristics

The countries included in this study were Australia, Finland, France, Norway, and the United States of America (USA) (Table 3.1). Data were publicly available from Finland's National Institute for Health and Welfare (THL) and France's Centre National de References des Pneumocoques (CNRP). Data from Norway and Australia were requested from Meldesystem

for Smittsomme Sykdommer (MSIS) and National Notifiable Disease Surveillance System
(NNDSS) respectively. Data from the USA were obtained from a published article [322].

The breakdown by age groups was the same in Australia, Finland and Norway, and similar
to the USA. France had only two age groups, making a single cut off for children $\leq$ 16 years
and adults $>$ 16 years. As such, I grouped older adolescents with children in all countries
to maintain consistency, such that all individuals $<$ 18 years were considered children and
all those $\geq$ 18 years were considered adults. France included data pre-PCV7 as well as
pre- and post-PCV13, whereas Australia, Norway and USA only had pre- and post-PCV13
data available. The first two years of Australian data were removed as they had limited
serotype data. The USA data taken from the published article included only the five most
common NVTs [322]. Finland, which implemented PCV10, had pre- and post-PCV10 data
available. Because PCV7 was implemented only one year prior to PCV10 implementation,
the Finnish pre-PCV10 data included the PCV7 implementation year. Most countries had
a 2+1 vaccine schedule, apart from Australia (3+0) and the USA (3+1). While PCV7 was
implemented in various years, the higher valent vaccines (PCV10/13) were implemented
in 2010 or 2011 for all countries. Overall national IPD incidence decreased in all countries
following the introduction of PCV7 and PCV10/13, with the biggest decline in the USA after
PCV7, followed by the introduction of PCV13 in Norway (Figure 3.1).

Table 3.1: Vaccination details for Australia, Finland, France, Norway and the United States, including the years of introduction, dosing schedule, surveillance data source, years of vaccination data available, and the serotyping performed. Abbreviations: PCV: pneumococcal conjugate vaccine.

	PCV Introduction	Vaccine doses	Surveillance Data Source	Years of vaccination data available	Serotyping
Australia	PCV7: 2005 PCV10: 2009 (some jurisdictions) PCV13: 2011	3+0	National Notifiable Disease Surveillance System (NNDSS)	Pre-PCV7: 1999-2005 Pre-PCV13: 2006-2011 Post: 2012-2016	Quellung reaction, molecular serotyping
Finland	PCV7: 2009 PCV10: 2010	2+1	National Institute for Health and Welfare (THL)	Pre-PCV10: 2004-2010 Post: 2011-2016	Quellung reaction
France	PCV7: 2006 PCV13: 2010	2+1	Centre National de References des Pneumocoques (CNRP)	Pre-PCV7: 2001-2006 Pre-PCV13: 2007-2010 Post: 2011-2016	Fourier Transformation-Infrared Spectroscopy, Multi Locus Sequence Typing, genomic sequencing
Norway	PCV7: 2006 PCV13: 2011	2+1	Meldesystem for Smittsomme Sykdommer (MSIS)	Pre-PCV13: 2006-2011 Post: 2012-2016	Quellung reaction
USA	PCV7: 2000 PCV13: 2010	3+1	CDC's Emerging Infections Program / Active Bacterial Core Surveillance	Pre-PCV13: 2005-2010 Post: 2011-2013	Quellung reaction
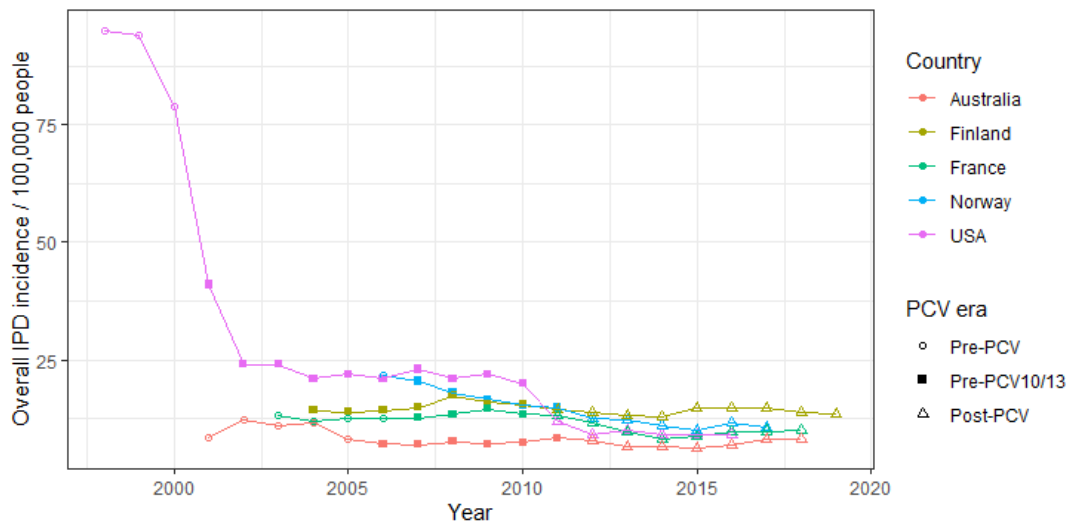
Figure 3.1: Incidence of IPD per 100,000 people in Australia, Finland, France, Norway and USA pre- and post-vaccination with PCV7 and PCV10/13 from national surveillance of all serotypes.

3.2.2   Temporal trends by serotype category

The age group-specific odds ratio (OR) compared the difference in disease caused by one serotype category (for example, VT7) after vaccine introductions, compared to all other serotype categories and PCV eras, to determine the correlation between a vaccine employed and an increase or decrease in the proportion of disease caused by that serotype category. This shows the trends in relative incidence for different serotype categories, avoiding fluctuations in absolute prevalence that may result from non-vaccine factors. According to the OR, vaccines decreased IPD caused by the serotypes they targeted in infants in all countries, with the exception of VT10 types in one age category in Finland (Figure 3.2). Decreases in VT7 IPD in all adult age categories implied a strong herd immunity effect. PCV13 consistently reduced IPD caused by VT13 in infants across countries where it was introduced, but the impact in adults was small relative to PCV7. For example, in the US and Norway, young children under 5 and the elderly were significantly less likely to get disease from VT10 and VT13 than other age groups post-PCV13, but other age groups had no significant association between PCV13 and disease caused by VT10 and VT13. The proportion of IPD caused by NVT significantly increased across all age groups and vaccination periods in Australia, France, Norway and the

US, even in categories in which no significant decrease in vaccine types had been detected.
There is evidence of herd immunity in adults caused by PCV10 administration to infants,
with accompanying serotype replacement by non-VT10 serotypes across all ages. This is
due to serotypes not included in PCV10 (VT13 and NVT) causing a greater proportion of
disease in all age groups while non-vaccinated age groups are benefiting from vaccination
with reduced proportion of disease caused by VT7 and VT10. Comparing across age groups
shows herd immunity effects, despite very different absolute incidences in these different
demographics. This increase in non-VT OR may represent the simple removal of VTs, or
expansion of NVTs in a serotype replacement process. These two possibilities can be resolved
through further analysis of IPD trends.
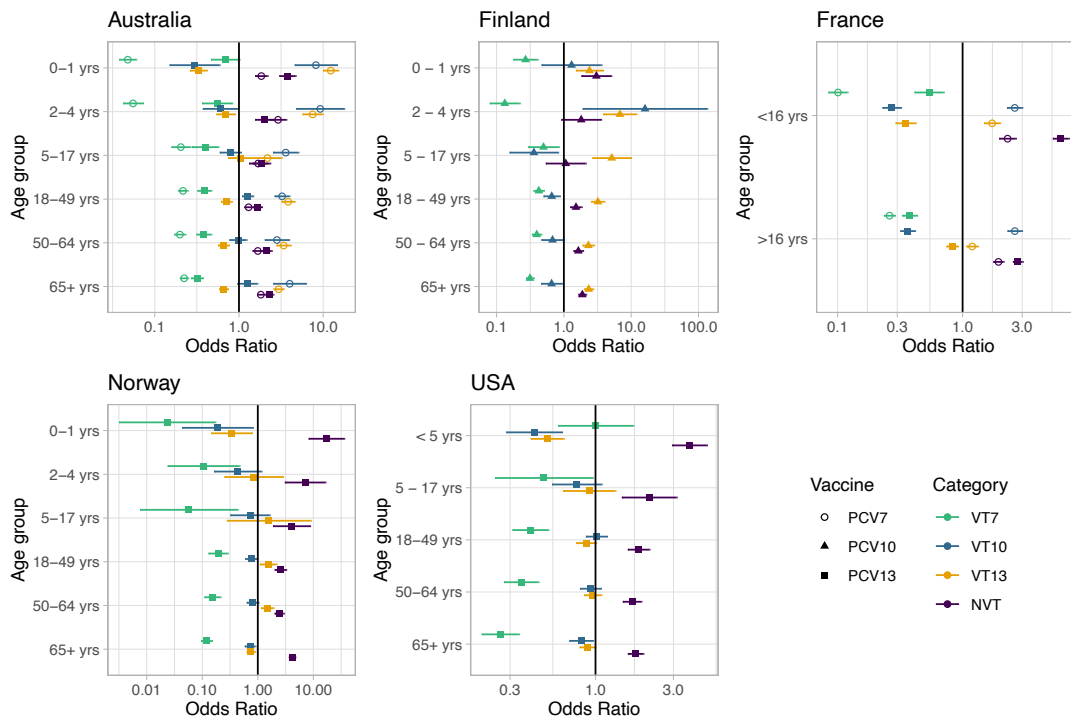
Figure 3.2: Point estimates and 95% confidence intervals (CI) of age-stratified odds ratios
(OR) relating IPD caused by serotype category after vaccine implementation. Estimated as
the product of the IPD cases by a serotype category post-vaccination and the IPD cases by all
other serotype categories pre-vaccination divided by the product of IPD cases by the serotype
category pre-vaccination and the IPD cases by all other serotype categories post-vaccination.
VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.

3.2.3 Effect of vaccination on serotype diversity

Simpson's Diversity Index (SDI) is a measure of diversity that assesses the abundance of species in a population. In this case, estimating SDI in a population before and after vaccination can help elucidate whether vaccination changes the distribution of the serotypes causing diseases. Populations with a few dominating serotypes have a lower diversity than populations with a more even distribution of serotypes. While the SDI in adults remained relatively stable through the vaccination phases, indicating the population structure of serotypes causing disease stays constant, significant changes were observed in children (Figure 3.3). Prior to vaccination with PCV7, adult IPD isolates had a significantly higher SDI than those from children in Australia, Finland and France, implying a greater variety of serotypes caused adult IPD, which is in agreement with previous published research [136]. In Australia, France and the US, there was a fall in SDI between the introduction of PCV7 and PCV13, consistent with a few VT13 causing a large proportion of IPD. After PCV13 introduction, the elimination of these VT13 serotypes meant the SDI of disease in children increased significantly and was no longer significantly different from that of adults. This elevated SDI suggests reducing infant IPD further may require targeting a wide diversity of serotypes.

Data from Australia, France and Norway were sufficiently comparable for them to be pooled over the first 4 years post-vaccination with PCV7 and PCV13 for children and adults respectively. This enables SDI to be estimated internationally. Finland and the USA were excluded from this analysis in order to maintain consistency between the countries being compared; Finland was excluded because it did not implement PCV13 and the window of PCV7 implementation was too narrow (only one year before PCV10 was implemented); and the USA was excluded because the data available did not include immediate post-PCV7 data. Mirroring the within-country results, there was a notable divergence between children and adults in the two post-vaccination periods (Figure 3.4). SDI decreased significantly post-PCV7 in children but did not change significantly in adults, suggesting the same serotypes were dominating post-PCV7 across countries. However, post-PCV13 SDI increased dramatically in children

but slowly in adults. After PCV7, diversity was considerably different between children and
adults whereas post-PCV13, the diversity was not significantly different between the two age
groups. This suggests the diversity of serotypes causing child disease is now similar to that
in adults, both within and between countries, making it more difficult to target with limited
valency vaccines.
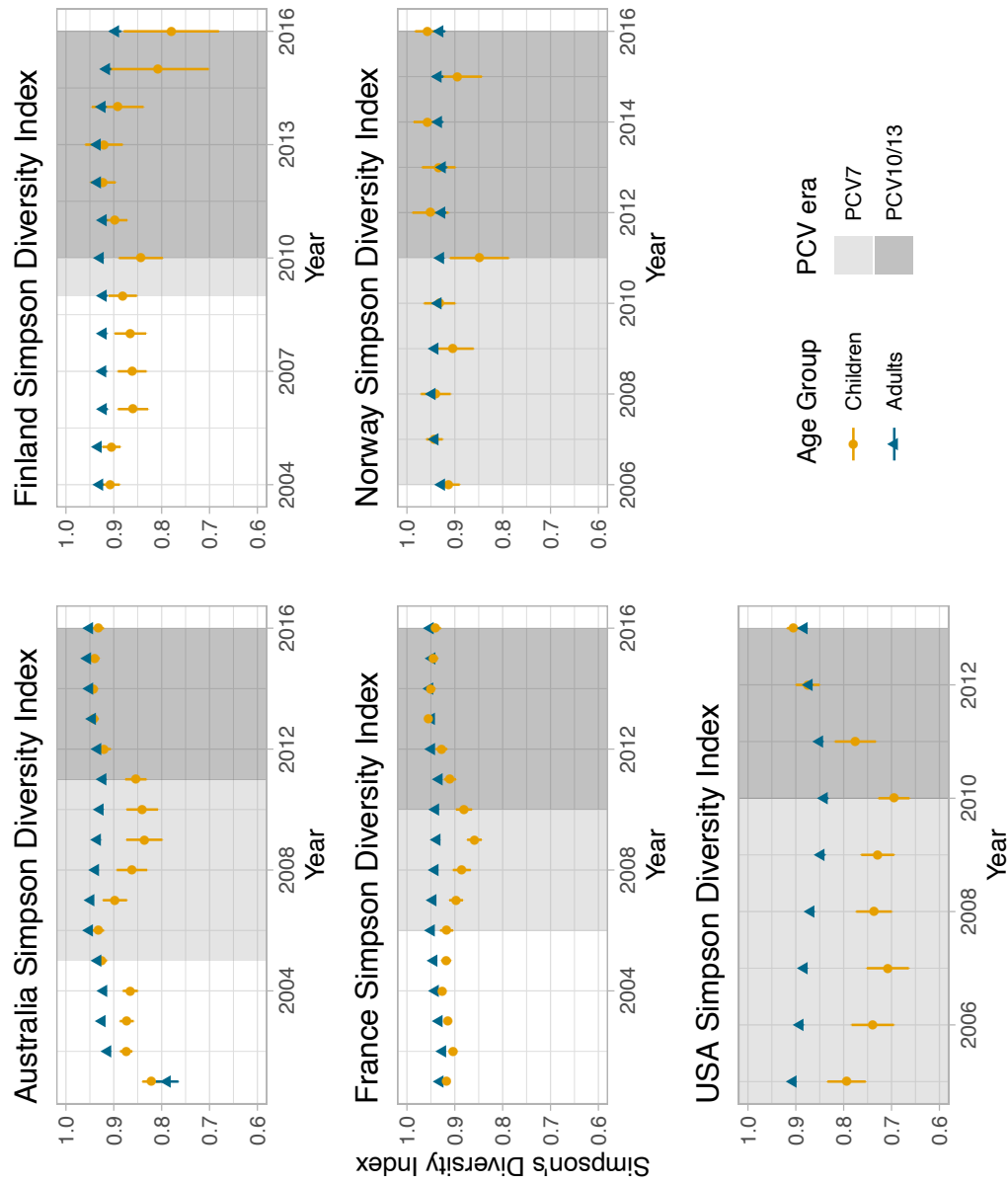
Figure 3.3: Simpson Diversity Index point estimate and 95% confidence intervals of serotypes causing IPD pre- and post-vaccination in children and adults. Children: ≤ 16 years for France, < 18 for all other countries; Adults: > 16 years for France, ≥ 18 years for all other countries.
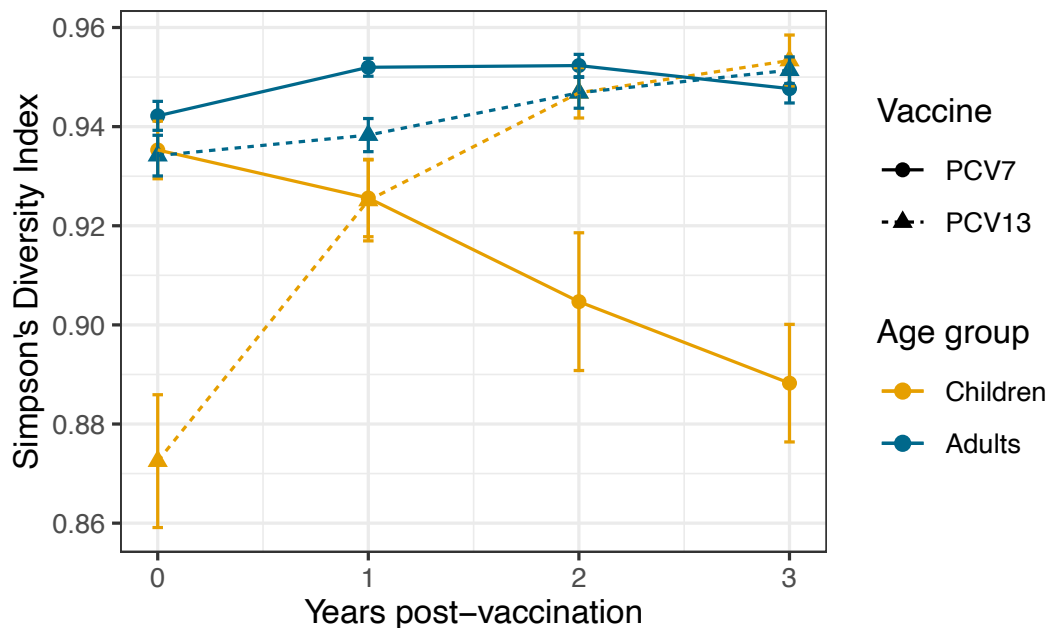
Figure 3.4: Simpson's Diversity Index three years post-vaccination with 95% confidence intervals pooling Australia, France and Norway in children ($\leq$ 16 years for France, $<$ 18 years for Australia and Norway) and adults ($>$ 16 years for France, $\geq$ 18 years for Australia and Norway).

3.2.4 Comparison of countries' rank frequency distribution and major IPD-causing serotypes

To identify the serotypes underlying these trends, rank frequency plots showed the ranking of the top fifty disease-causing serotypes in each PCV era by the number of disease cases caused and a cumulative frequency curve of the total disease (Figures 3.5, 3.6, 3.7, 3.8, 3.9). After vaccination, VT IPD decreased, and NVT rose up in rankings in both age groups and countries. Each country and age group experienced, in respective pre-PCV eras, one to three serotypes causing a majority of the disease burden. Post-PCV, there were fewer 'dominating' serotypes, defined as those causing at least twice the IPD cases as that of the next-ranked serotype. This is also seen in the cumulative frequency curves, where the number of serotypes causing around 80% of disease in both age groups before vaccination was below ten, but increased after vaccination with PCV7 and PCV13. This number remained relatively stable following the introduction of the higher valency vaccines. The exception was in Norway,

90

where there were no dominating serotypes pre- or post-PCV13 in both age groups (Figure 3.8).
Serotype replacement is evident post-vaccination in all countries, which is seen independent
of national surveillance agency [323], except in the US, where the most common serotypes
are in the same serotype category (VT13 and VT10) after vaccination.

The rank frequency plots suggest the most important trends in IPD can be explained by
analysing the ten most prevalent serotypes. The top IPD-causing serotypes differed between
children and adults and diverged post-vaccination (Figure 3.10, Table 3.2). Pre-vaccination,
the countries had a similar composition of serotypes in their rankings which consisted
primarily of VT. Serotype 14 was often dominant in both child and adult IPD. Post-vaccination
with PCV7 (i.e. pre-PCV10/13), serotypes 7F and 19A dominated rankings in children, as well
as serotype 3 in some settings, explaining the fall in post-PCV7 SDI. The rankings diverged
post-PCV13 between children and adults, and in the same age group between countries, with
NVT dominating the rankings. Adult IPD was caused by an overlapping set of NVT when
compared with child IPD post-vaccination (Table 3.2). While some serotypes like 12F and
23B seemed to affect both age groups, serotypes 8 and 9N tended to appear in adult rankings
more frequently and serotype 38 was more common in children. VT10 and VT13 serotypes
3, 7F and 19A persisted post-PCV10/13 in adults, with serotype 3 actually rising up the
rankings in some locations. Serotype 19A still dominated the infant rankings post-PCV13
in many countries except Norway. In line with previously reported analyses [324], there
was no dominating NVT that emerged post-PCV in the same fashion that serotype 19A did
post-PCV7. With the exception of the USA, the rankings were not consistent with 'unmasking'
of NVT [175], as NVT changed in their relative prevalence following PCV introduction. This
suggests a serotype replacement process driven by the post-PCV expansion of certain NVTs.
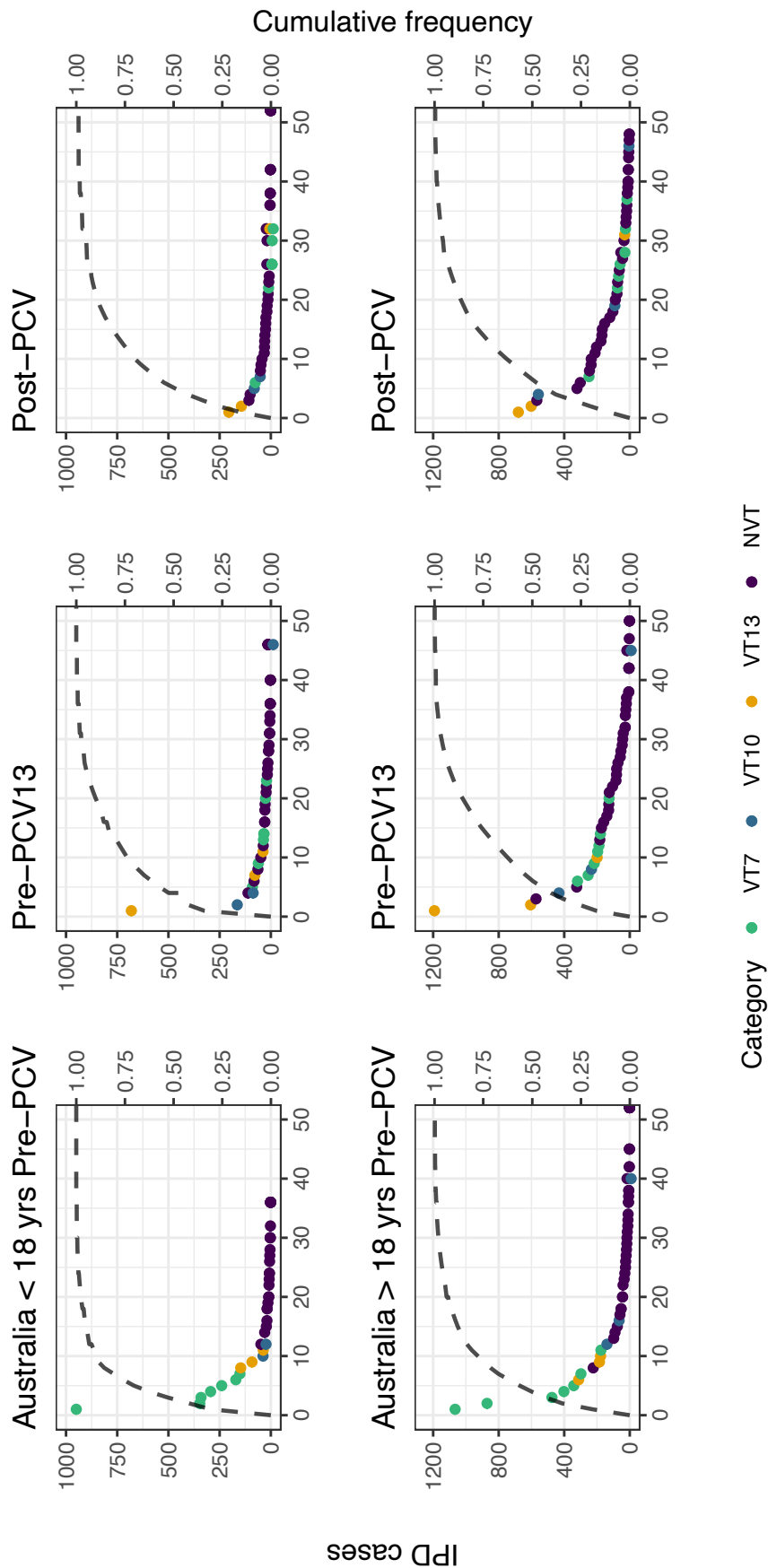
Figure 3.5: Rank frequency distribution and cumulative frequency in IPD by serotype category in Australia. VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.
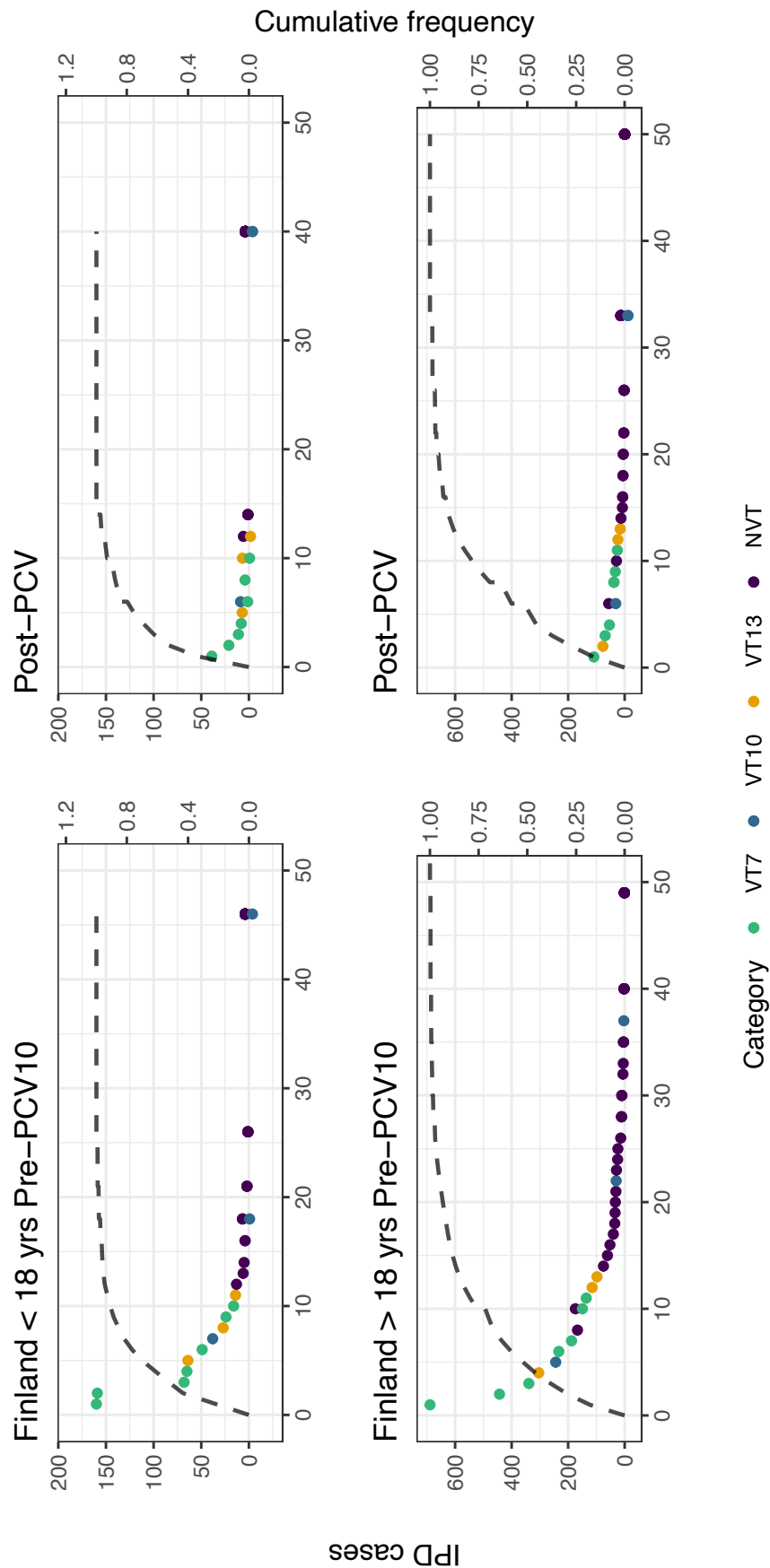
Figure 3.6: Rank frequency distribution and cumulative frequency in IPD by serotype category in Finland. VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.
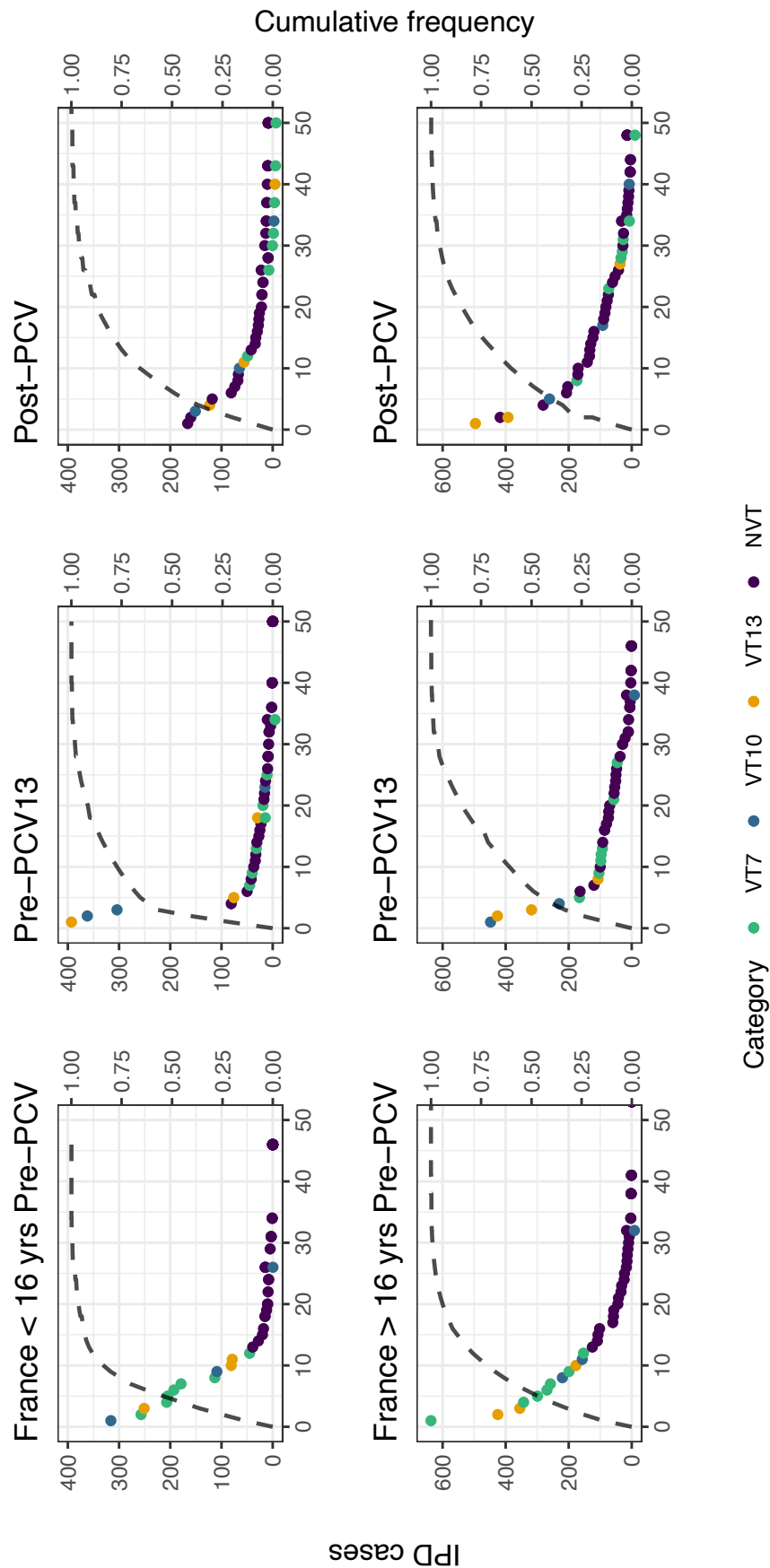
Figure 3.7: Rank frequency distribution and cumulative frequency in IPD by serotype category in France. VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.
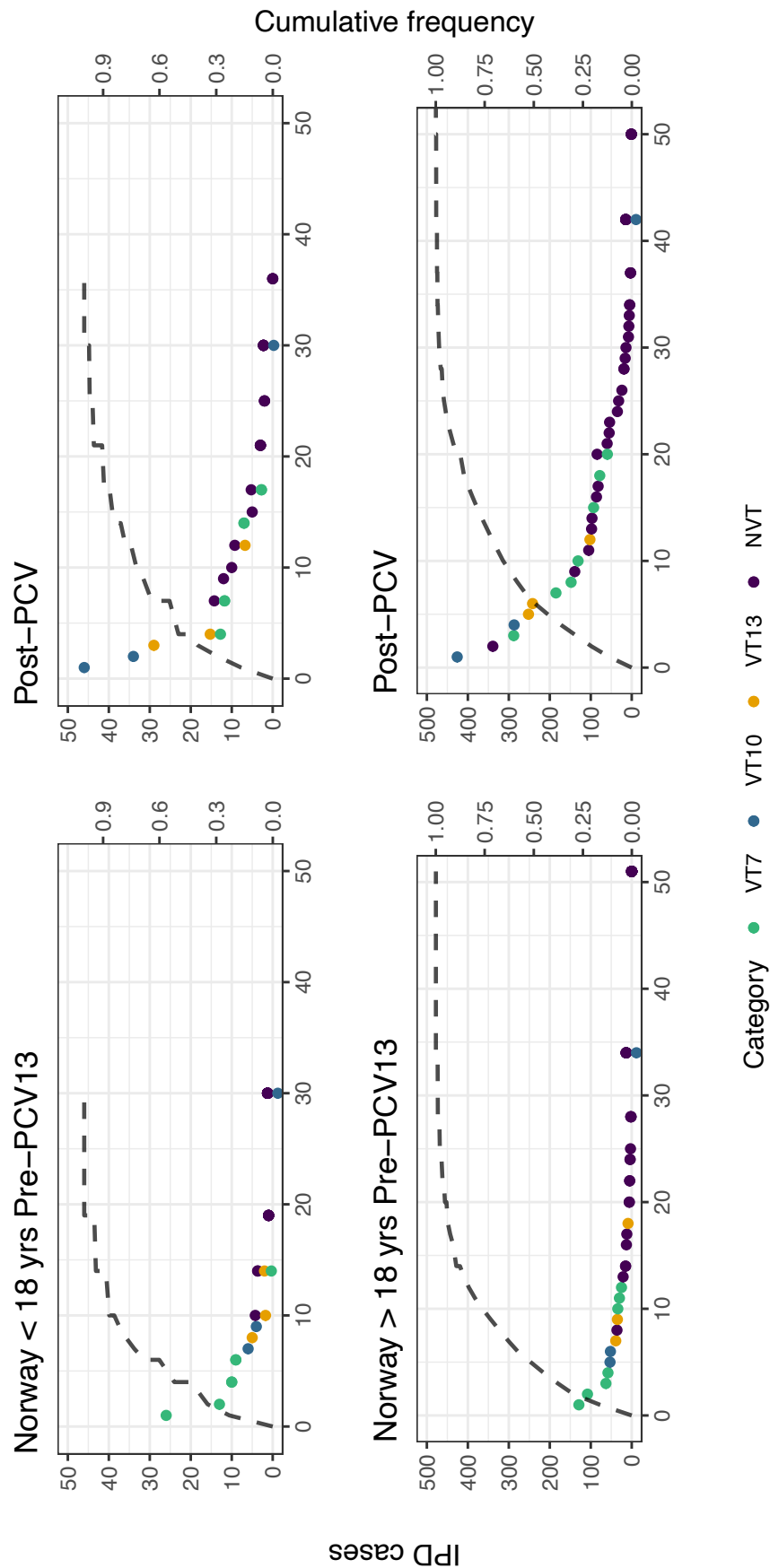
Figure 3.8: Rank frequency distribution and cumulative frequency in IPD by serotype category in Norway. VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.
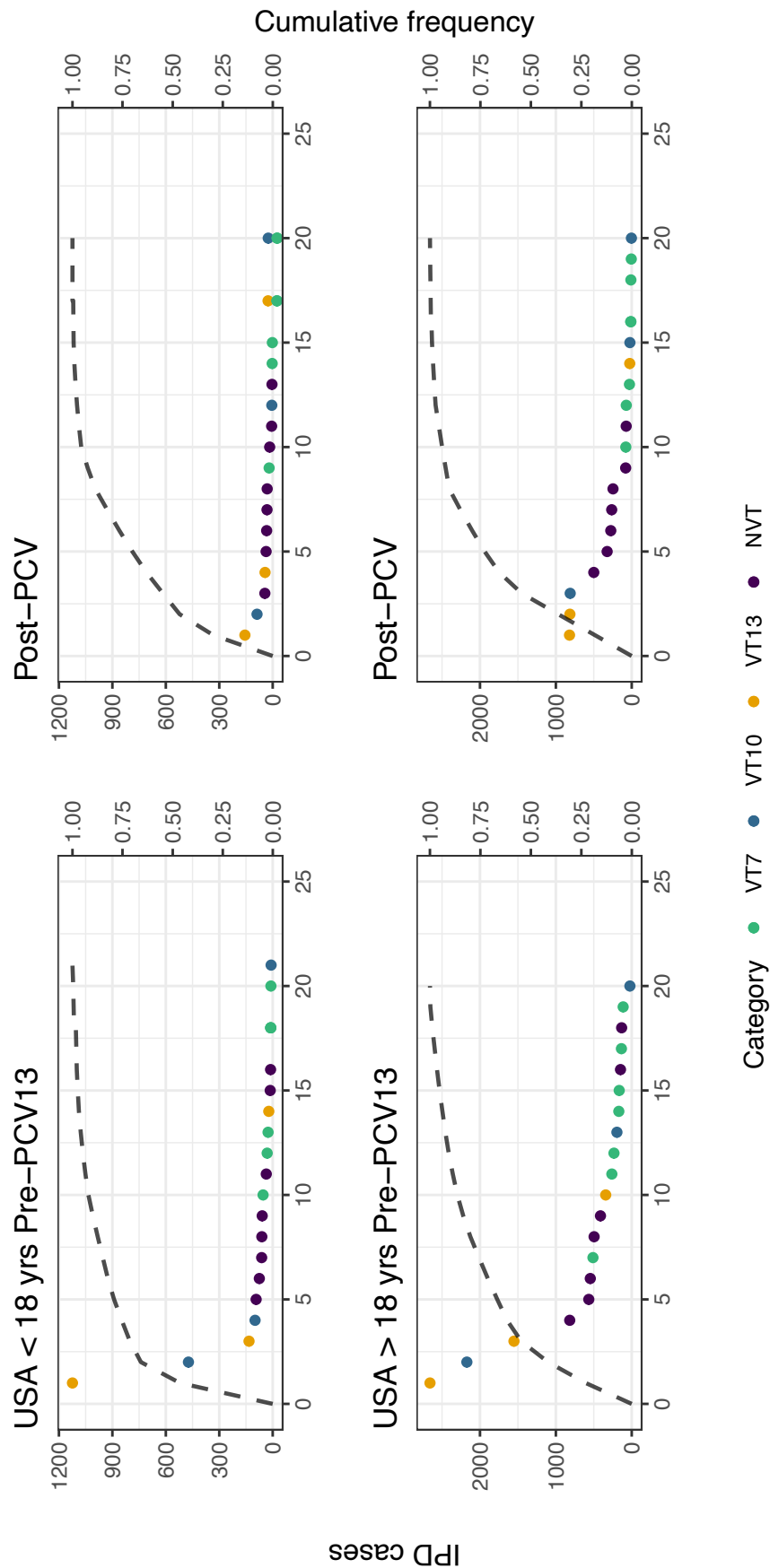
Figure 3.9: Rank frequency distribution and cumulative frequency in IPD by serotype category in USA. VT7: PCV7 serotypes; VT10: 1, 5, 7F; VT13: 3, 6A, 19A; NVT: non-vaccine types.
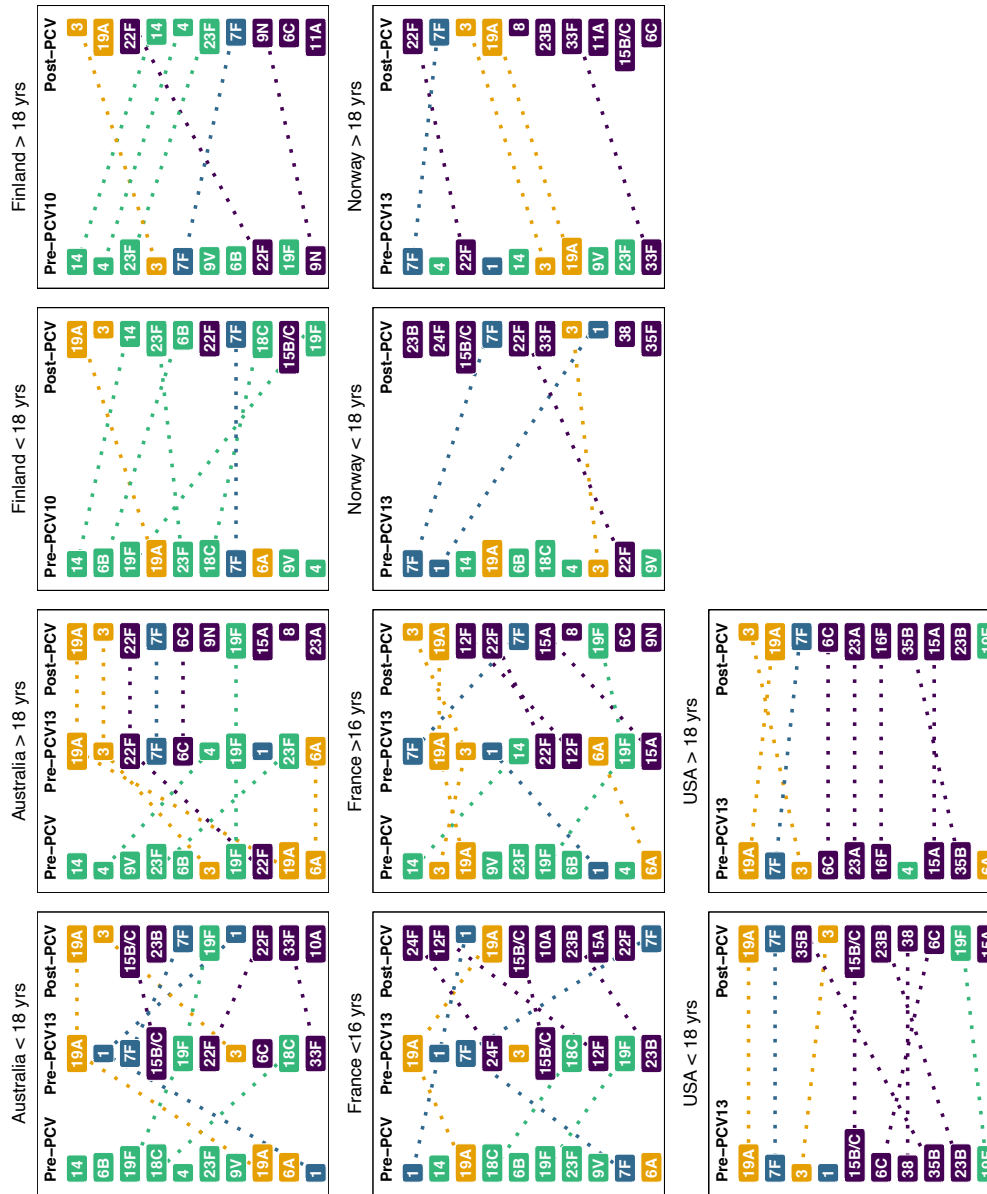
Figure 3.10: Top ten IPD-causing serotypes pre- and post-vaccination in (A) Australia (B) Finland, (C) France, (D) Norway and (E) USA, with colour corresponding to serotype category. Green: VT7 (PCV7 serotypes); blue: VT10 (1, 5, 7F); yellow: VT13 (3, 6A, 19A); purple: NVT.

Table 3.2: Aggregated top serotypes causing IPD in Australia, Finland, France, Norway and the United States by age group and PCV era; Colours correspond to serotype category. Green: VT7; blue: VT10; yellow: VT13; purple: NVT. Serotypes with an asterisk indicated serotypes that do not occur in both age groups for that vaccination era.

	Pre-PCV7	Pre-PCV10/13	Post-PCV10/13
Children	4, 6B, 9V, 14, 18C, 19F, 23F, 1, 7F, 6A, 19A	4, 6B*, 9V, 14, 18C*, 19F, 23F, 1, 7F, 3, 6A, 19A, 6C, 12F, 15C*, 22F, 23B*, 24F*, 33F, 35B, 38*	6B, 14, 18C, 19F, 23F, 1, 7F, 3, 19A, 6C, 8, 10A*, 12F, 15A, 15B/C, 22F, 23B, 24F*, 33F, 35B, 35F*, 38*
Adults	4, 6B, 9V, 14, 19F, 23F, 1, 7F, 3, 6A, 19A, 22F	4, 9V, 14, 19F, 23F, 1, 7F, 3, 6A, 19A, 6C, 9N*, 12F, 15A*, 16F*, 22F, 23A*, 33F, 35B	14, 19F, 23F, 7F, 3, 19A, 6C, 8, 9N*, 11A*, 12F, 15A, 15B/C, 16F*, 22F, 23A*, 23B, 33F, 35B

The fraction of total IPD caused by all these major serotypes individually over time demonstrated a similar effect (Figures 3.11, 3.12, 3.13, 3.14). After any vaccine introduction, there was a steady decline in the frequency of VT7, stabilizing at levels below 5% of all IPD in most countries and all age groups. VT10 follow similar trends, although the decline in serotype 7F is still tapering off post-2012. Results are varied across settings for VT13 and NVT. VT13 serotype 6A, which was already affected by some cross-immunity with the PCV7 6B antigen, has decreased in adults and is consistently low in children. Across countries, there was a notable spike in the percentage of IPD incidence caused by serotypes 1, 7F and 19A between the introduction of PCV7 and PCV13. However, the serotype trends diverged between countries following the introduction of the expanded-valency PCVs. In the NVT serotype category, serotype 22F caused over 15% of disease in Norwegian adults post-PCV13 whereas it stayed stable around 7% in French adults. In children, serotype 22F has not increased dramatically and fluctuated greatly post-PCV10/13 in most countries. Serotype 24F caused over 20% of disease in Norwegian children in 2015 and over 15% in French children in 2016 (as in previous results [325]) but less than 3% in Finnish children during that same time. The fraction of disease caused by serotypes 6C and 35B in American adults increased sharply post-PCV13, whereas in other countries it stayed relatively stable between 2 and 6%. The percentage of serotype 8 increased gradually in most countries' adults, with a particularly dramatic rise in French adults.
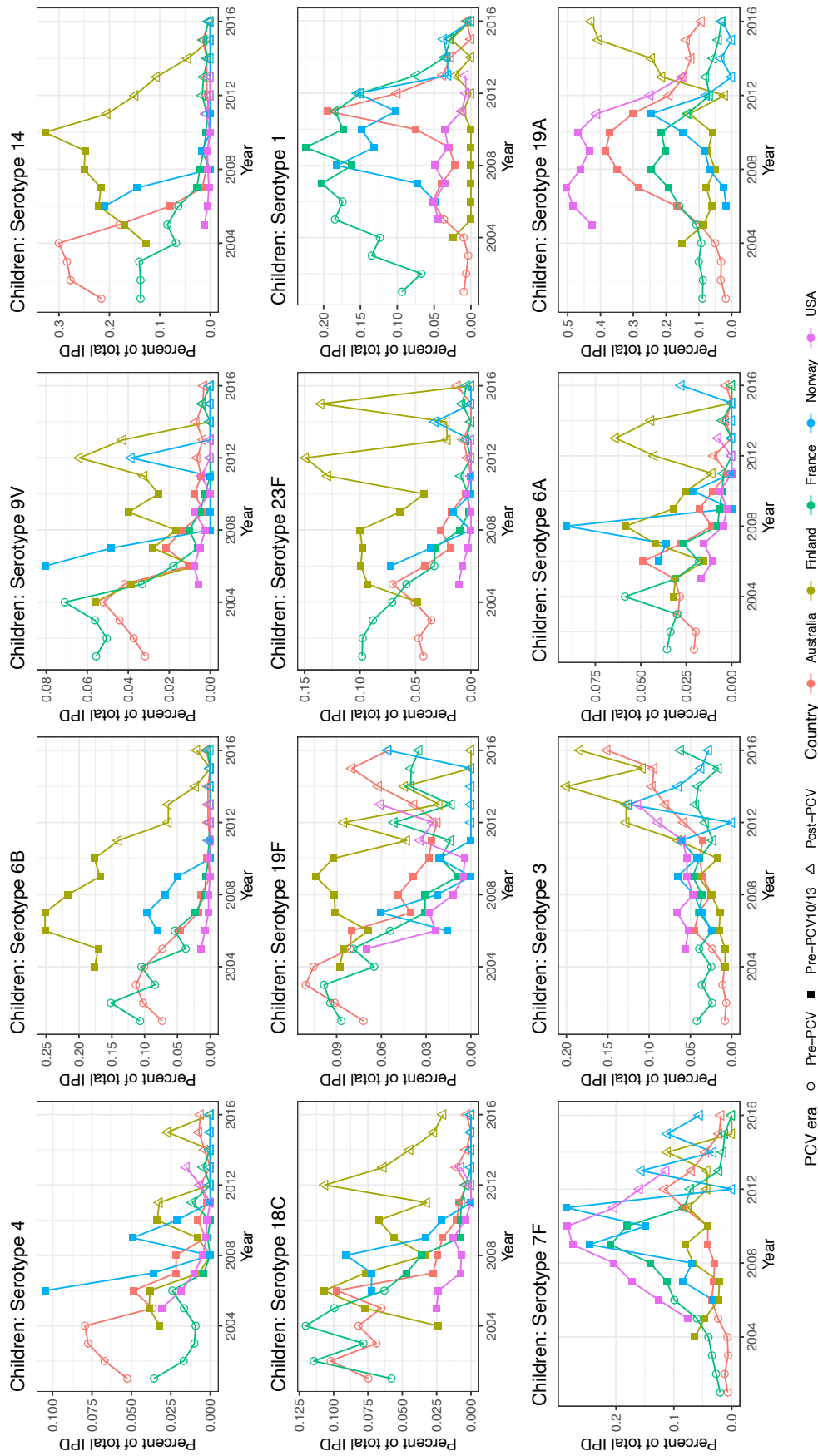
Figure 3.11: Change in percentage of national IPD caused by VT in children ($\leq$ 16 years for France, < 18 years for all other countries).
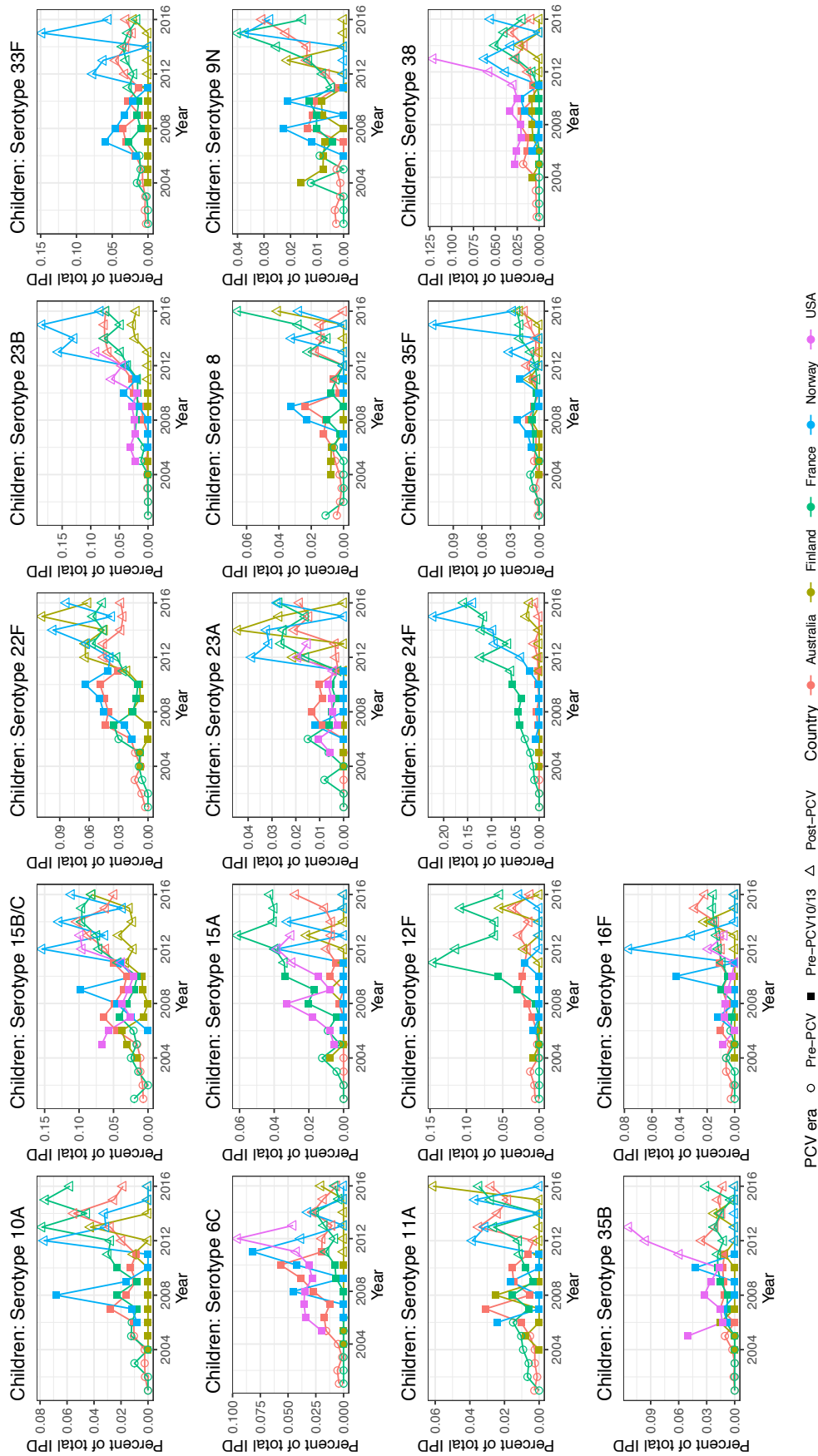
Figure 3.12: Change in percentage of national IPD caused by NVT in children ($\leq$ 16 years for France, < 18 years for all other countries).
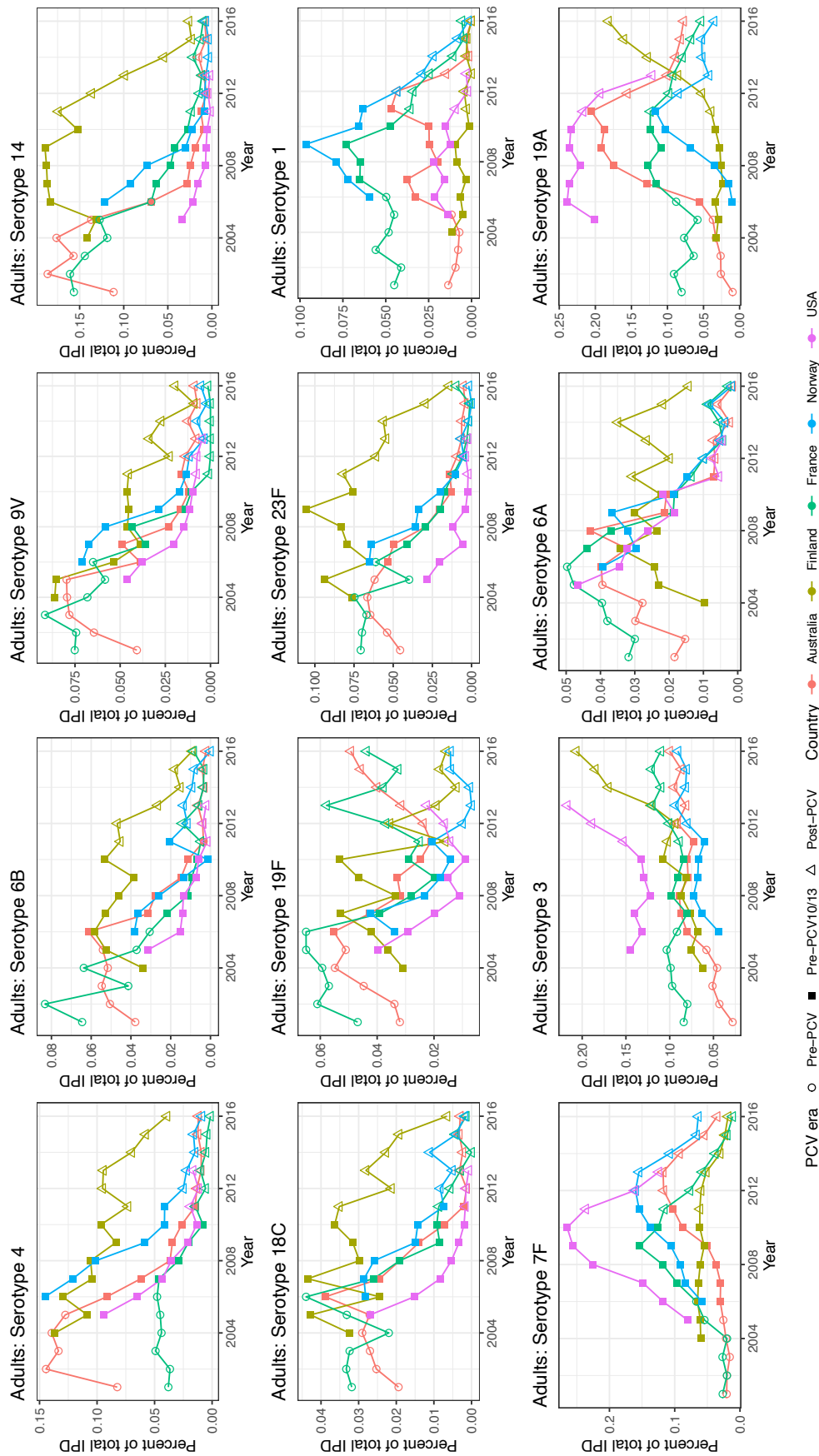
Figure 3.13: Change in percentage of national IPD caused by VT in adults ($>$ 16 years for France, $\geq$ 18 years for all other countries).
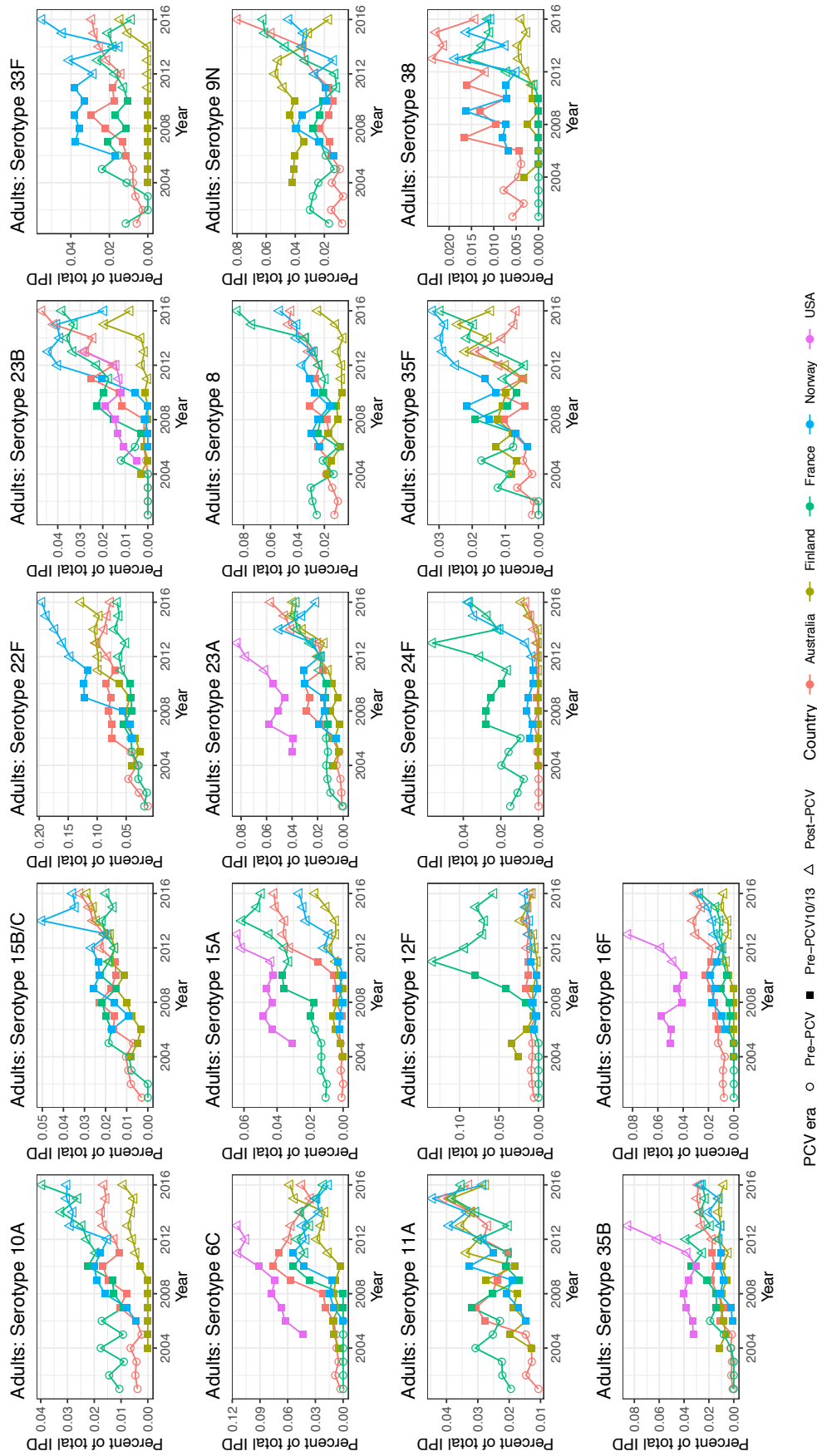
Figure 3.14: Change in percentage of national IPD caused by NVT in adults ($>$ 16 years for France, $\geq$ 18 years for all other countries).

3.2.5 Serotype incidence growth rate

Effects of vaccination can be seen in the rates of incidence change of certain serotypes across countries and age groups (Figures 3.15, 3.16, 3.17, 3.18 and Figures 3.19, 3.20, 3.21, 3.22). These highlight differences between countries in serotypes that are rare (absent from Figure 3.10) but growing due to serotype replacement, and serotypes that are common (present in Figure 3.10) but declining due to vaccination or herd immunity.

Australia

In Australia, VT7 were not significantly changing in children in incidence pre-PCV (except 18C) but were increasing significantly in adults (except 4 and 14). Post-PCV7, VT7 decreased in incidence, whereas VT10 and VT13 serotypes (except serotype 3 in children) were, as expected, increasing pre-PCV7 and PCV13 but decreasing post-PCV13 (Figure 3.23). Children and adults had similar incidence growth rates. Serotypes that were both common (Figure 3.10) and growing were 8, 9N and 22F (adults), and 23A and 23B (both age groups). Despite its persistent presence on the rank list post-vaccination, serotype 7F was decreasing substantially in both age groups post-PCV13, as was serotype 6C in adults. From these results, we see that post-PCV13 there has not been a sharp increase in any serotype as there was pre-PCV13 with serotype 19A.

Finland

Incidence of VT7 and VT10 in Finland decreased post-PCV10 in both age groups, whereas serotypes 3 and 19A (both VT13) grew sharply, particularly in adults (Figure 3.24). Serotype growth between children and adults tended to be similar, with differences being mostly non-significant. Serotype 19A decreased significantly in children compared to adults pre-PCV, whereas serotype 3 was growing more in adults than children. Despite being common in both age groups post-vaccination, serotype 14 decreased just as sharply post-vaccination as it increased pre-PCV in children and adults. Common serotypes that were increasing in rate of incidence were serotypes 6C and 22F in adults. On the other hand, serotype 8, 23A and

23B were all growing in incidence but rare in both age groups. This was the case for serotype
11A in children as well, although this serotype was already common in adult IPD. Finally,
despite 9N being a highly ranked serotype post-PCV in adults, the rate of incidence actually
declined.

France

The only VT serotypes increasing in incidence pre-PCV in France were serotypes 1 and 19A
(children) and 7F (both age groups), despite the rankings for both age groups consisting only
of VTs (Figure 3.10, Figure 3.25). Post-PCV13, common serotypes that were also increasing
significantly were serotypes 10A, 15A, 15B/C, 22F, 23B, and 24F in children and serotypes
8, 9N and 15A in adults. Common serotypes that were not increasing significantly were
serotypes 1 (children), 3, 6C, 19F, and 22F (adults), and 7F, 12F, and 19A (both age groups).
Serotypes that were rare (not included on the rank list, Figure 3.10) but increasing were
serotypes 8 and 9N (children), and 10A, 16F, 23B (adults) and 11A, 23A, 35F, 38 (both age
groups). Children and adults followed the same general trends of increase or decrease in
incidence.

Norway

Similar to other countries, VT7 serotypes were decreasing post-PCV7 in Norway in both
age groups (Figure 3.26). Serotypes 7F and 19A were the only VT10/13 serotypes increasing
after PCV7 in children and adults. Other serotypes increasing post-PCV7 were serotypes
24F (children), 10A, 22F and 23A (adults), and 6C and 23B (both age groups). Post-PCV13,
common serotypes increasing in incidence were serotypes 24F (children), 8 and 9N (adults),
and 22F (both age groups). Serotypes 1, 7F and 19A declined in incidence in adults, ex-
plaining the OR results from the previous section, and suggesting herd immunity. Common
serotypes not increasing or decreasing significantly were 3, 7F, 23B, and 33F in both age
groups. Surprisingly, serotypes 6A (VT13) and 19F (VT7) increased post-PCV in children.
Rare serotypes increasing in incidence but not common were serotypes 8, 9N, 12F and 35F

(children), and 10A, 12F, 15A, 16F, 24F and 35B (adults). Serotypes 8 and 9N were common in adults and 24F common in children.

USA

In the US, as expected, VT7 were all decreasing in incidence post-PCV7 in both age groups, whereas serotypes 7F and 19A were increasing (Figure 3.27). Other serotypes increasing significantly in incidence in this era were serotypes 15A, 23A and 23B (adults) and 6C (both age groups). Serotype 3 was the only serotype included in PCV13 that was increasing post-PCV13, although non-significantly. Other serotypes whose incidence was significantly increasing post-PCV13 were serotype 38 (children), 15A and 23B (adults) and 35B (both age groups). VT10 and VT13 serotypes 7F and 19A, despite being common in both age groups, decreased in incidence. Interestingly, serotypes 6A (VT13) and 18C (VT7) increased post-PCV in children, even though they were not common. The similarities between age groups suggests herd immunity from vaccination.
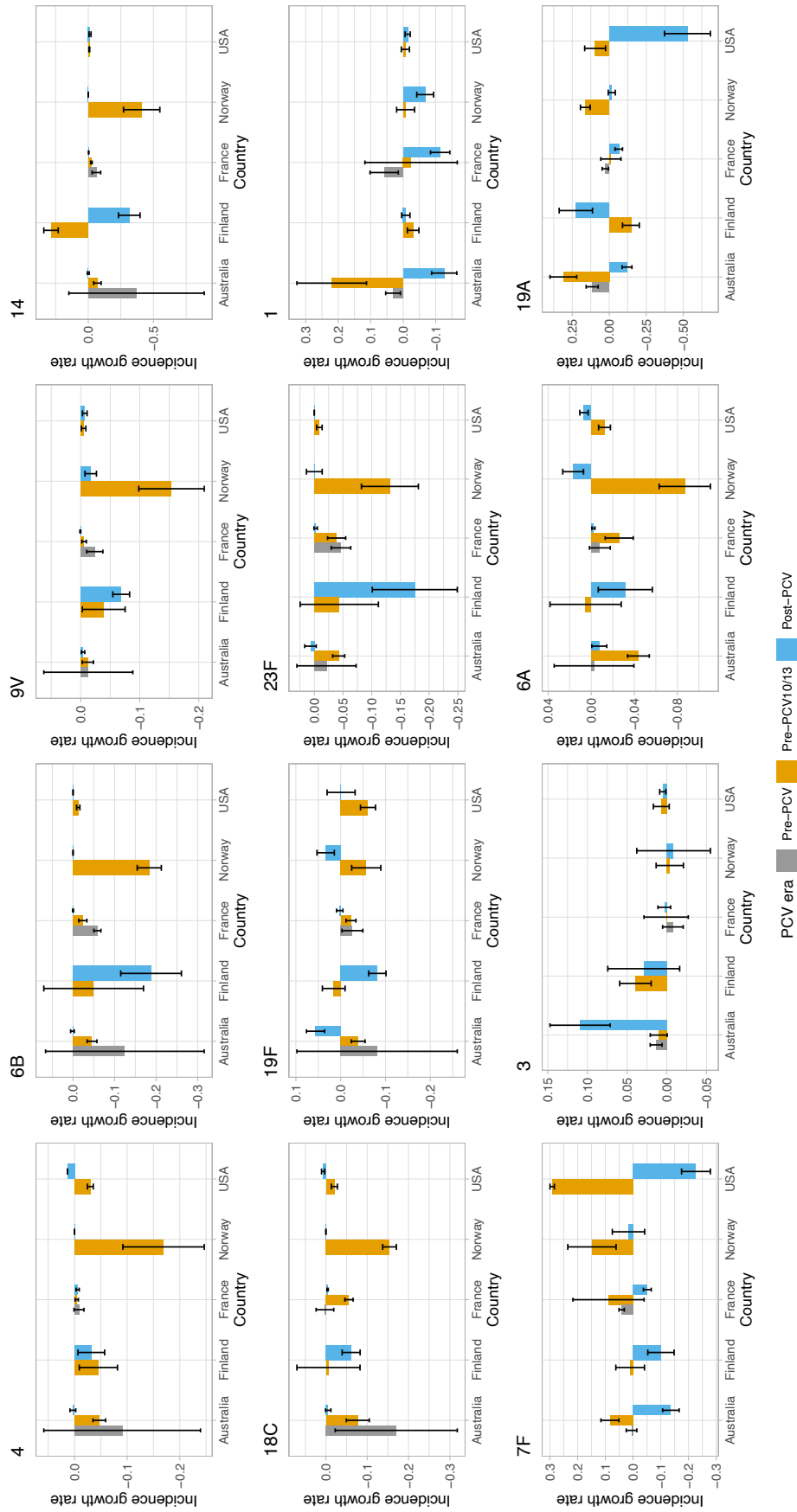
Figure 3.15: Comparison of incidence growth rates of VT in children between countries. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years).
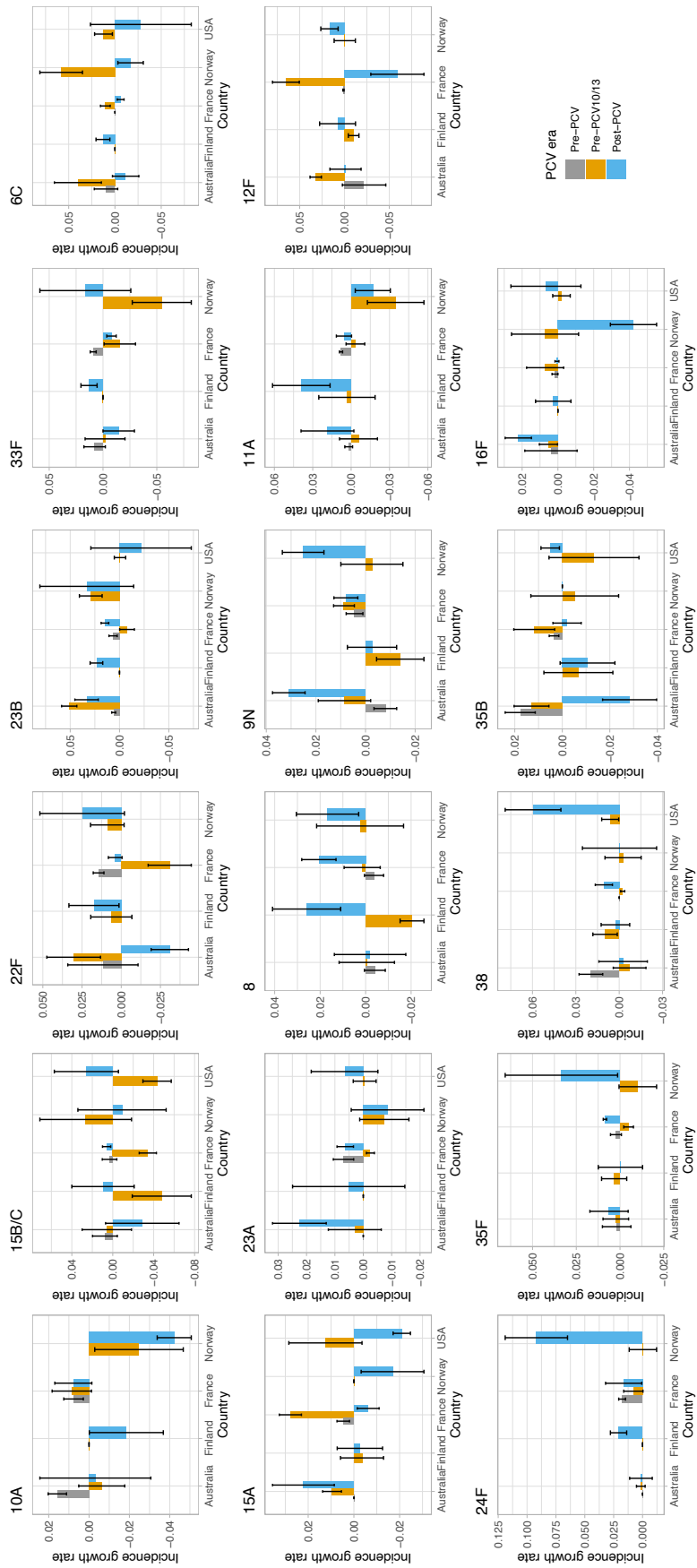
Figure 3.16: Comparison of incidence growth rates of NVT in children between countries. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years).
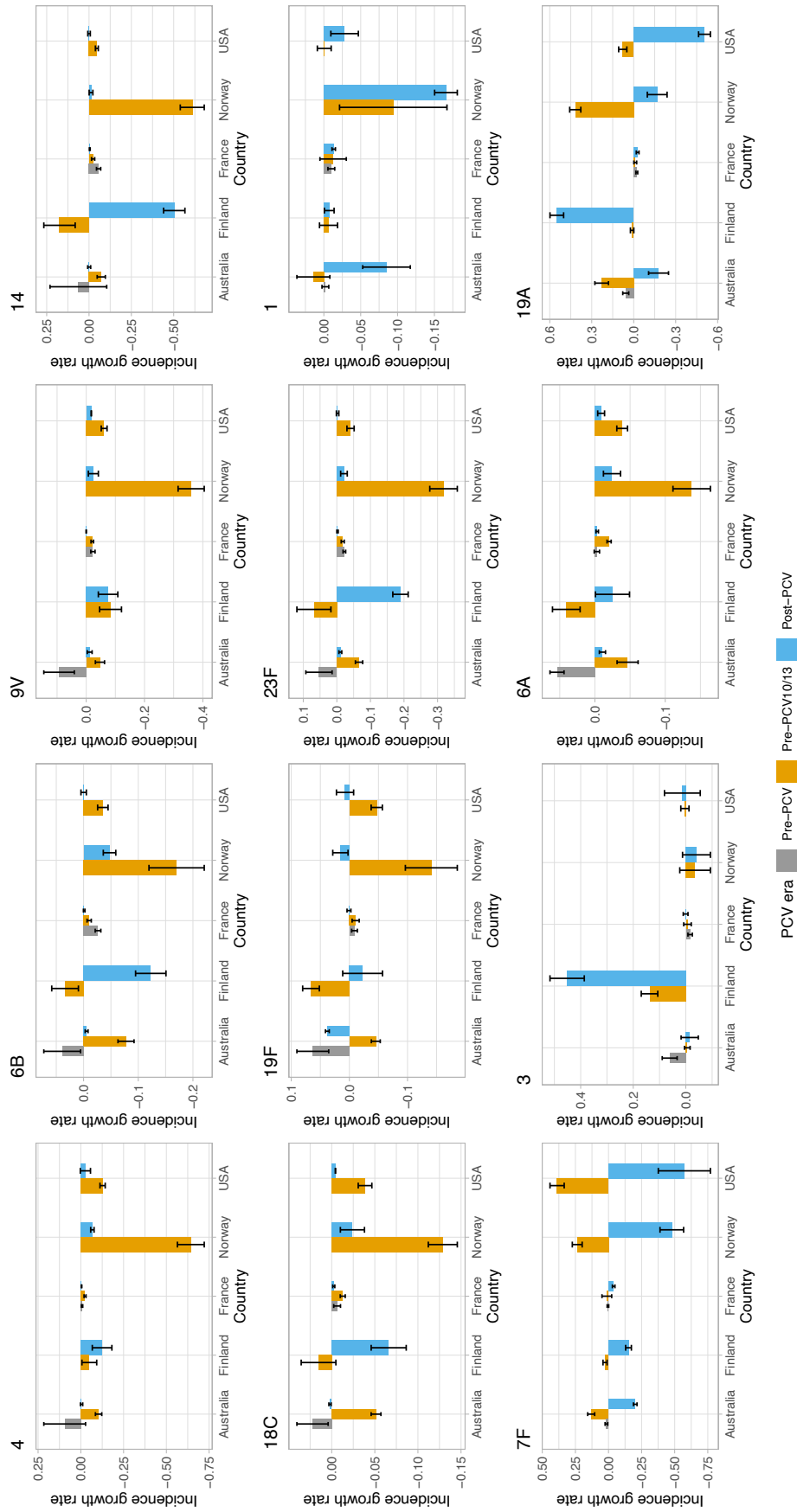
Figure 3.17: Comparison of incidence growth rates of VT in adults between countries. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years).
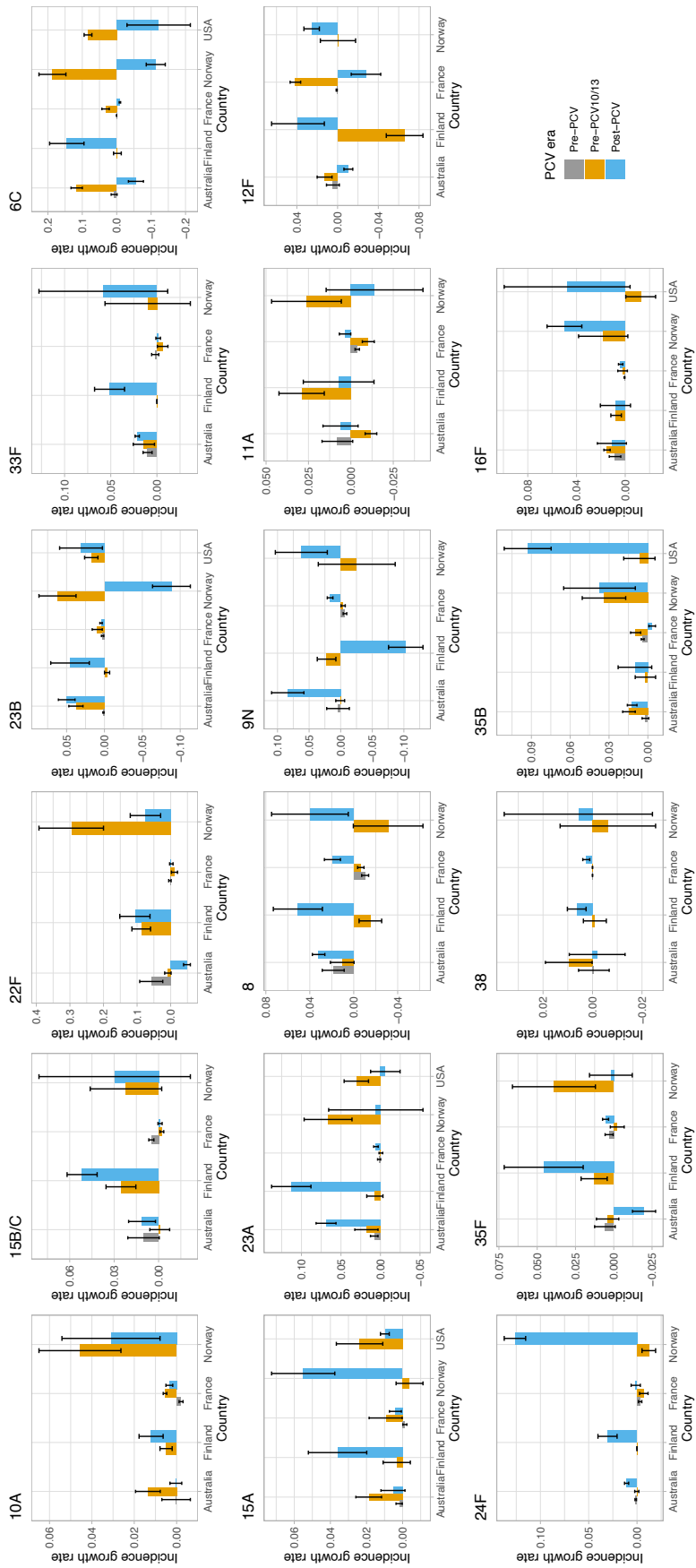
Figure 3.18: Comparison of incidence growth rates of NVT in adults between countries. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years).
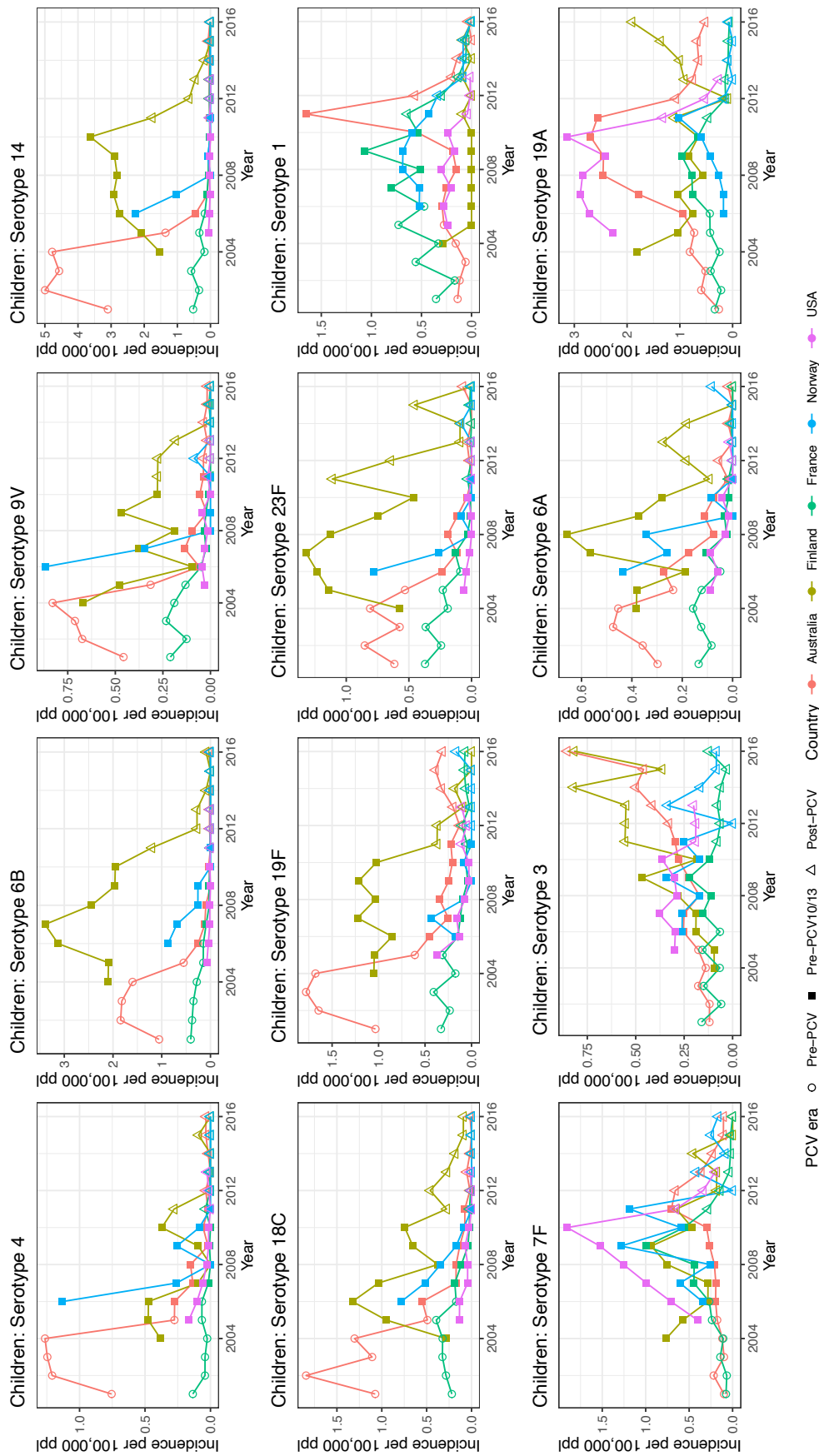
Figure 3.19: Incidence of VTs in children in different countries.
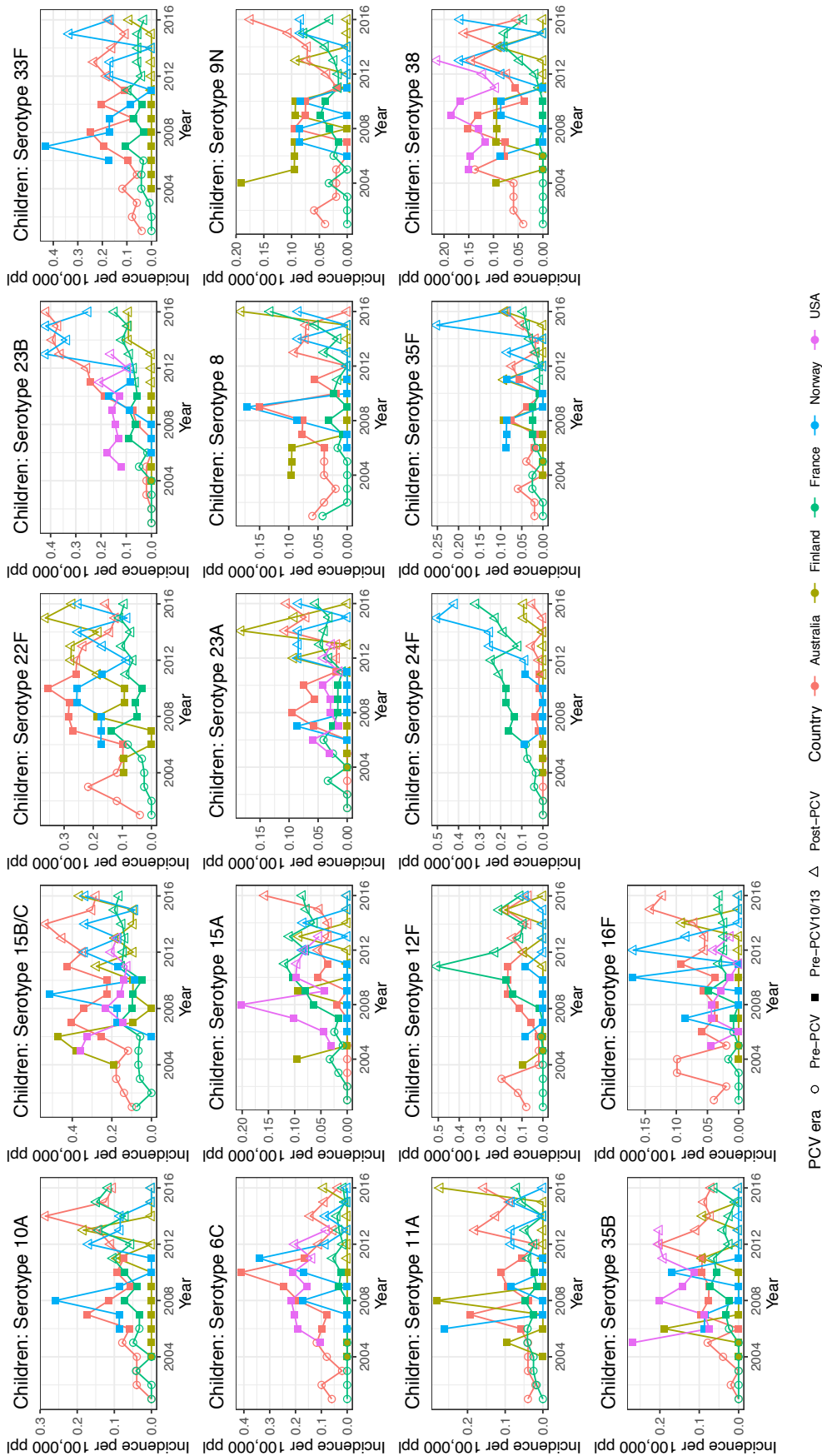
Figure 3.20: Incidence of NVTs in children in different countries.
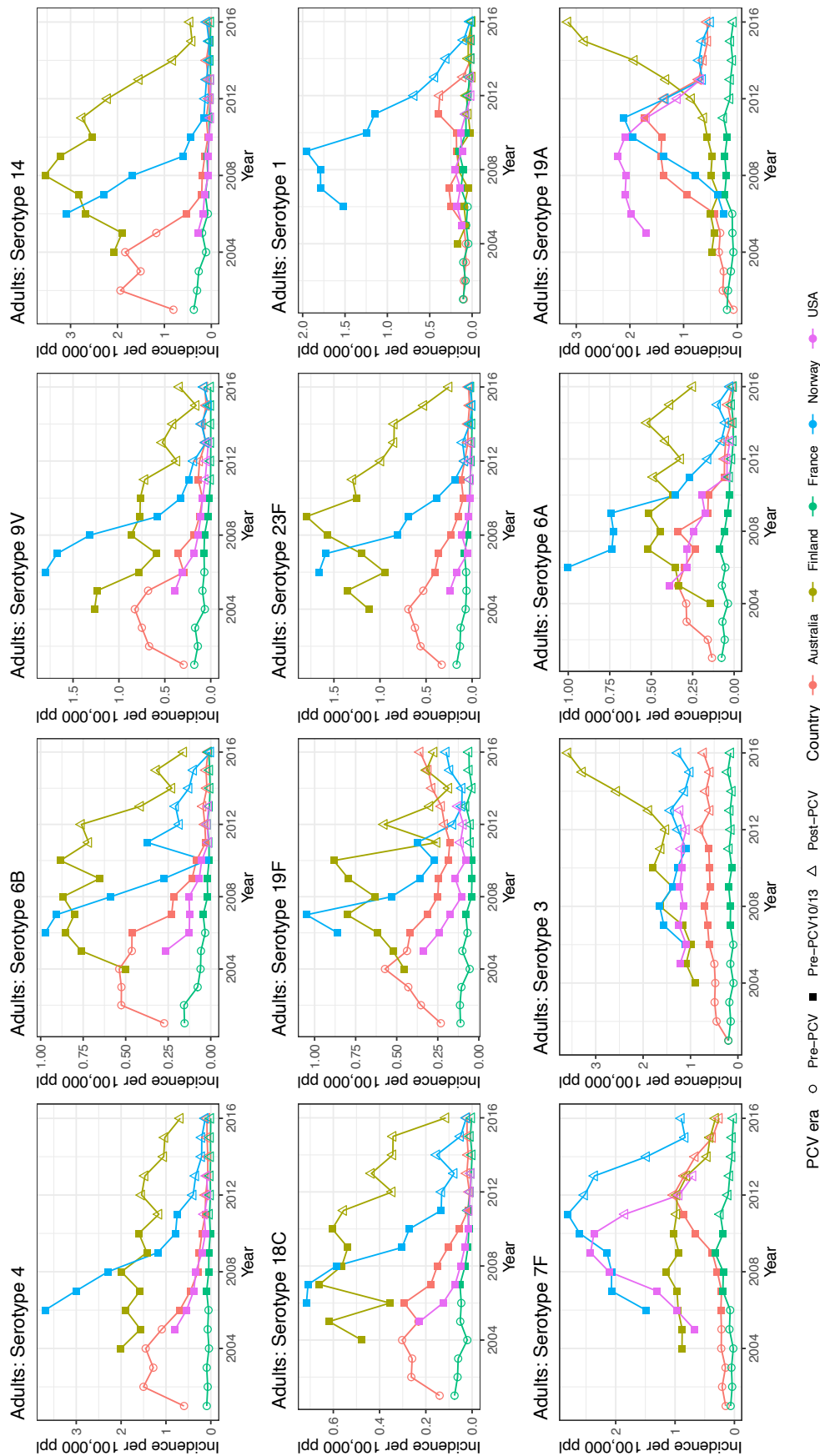
Figure 3.21: Incidence of VTs in adults in different countries.
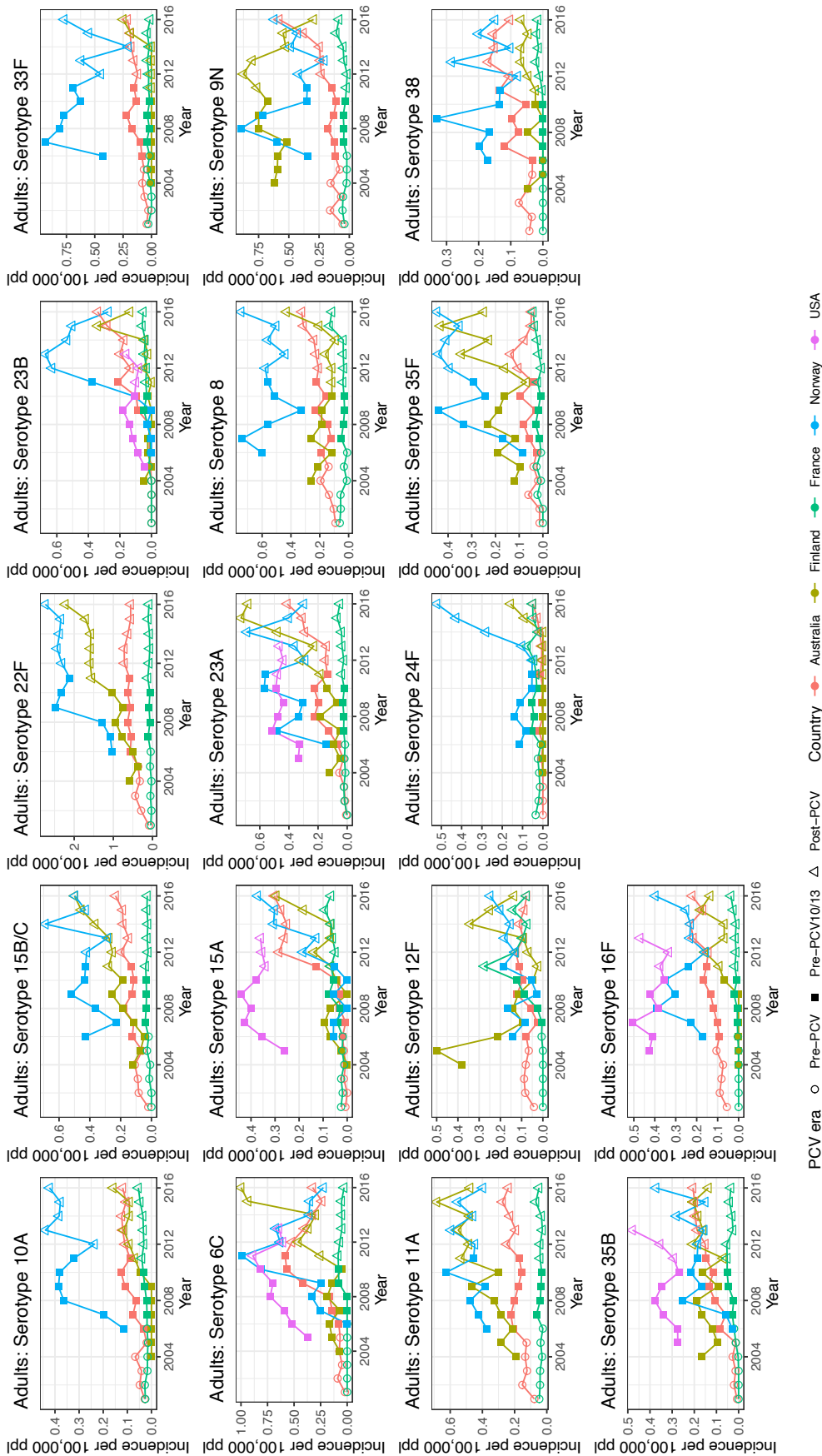
Figure 3.22: Incidence of NVTs in adults in different countries.
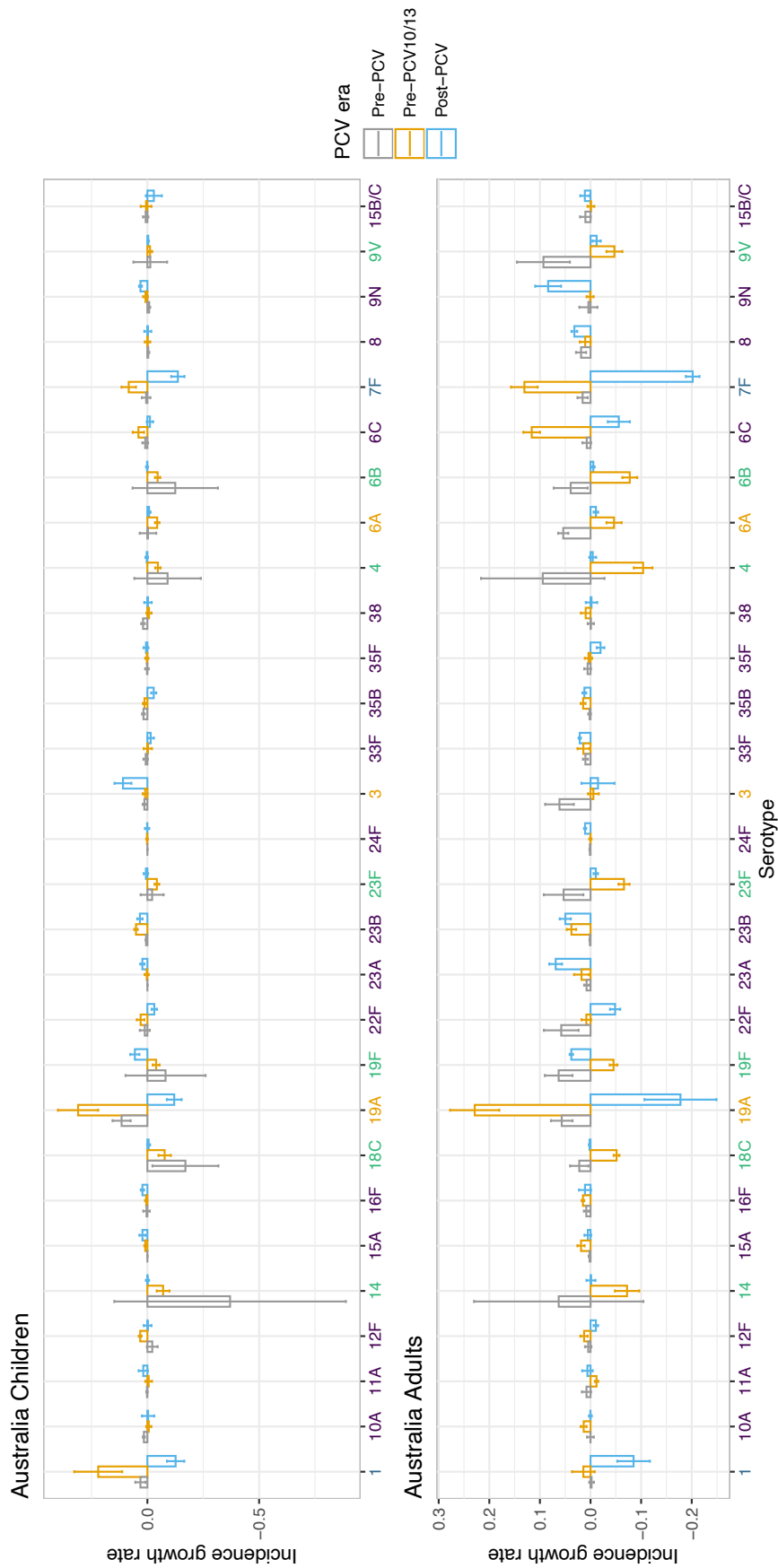
Figure 3.23: Rate of incidence growth in Australia IPD-causing serotypes pre- and post-vaccination. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years). X-axis labels in green: VT7 (PCV7 serotypes); x-axis labels in blue: VT10 (1, 5, 7F); x-axis labels in yellow (3, 6A, 19A); x-axis labels in purple: NVT (non-vaccine types)
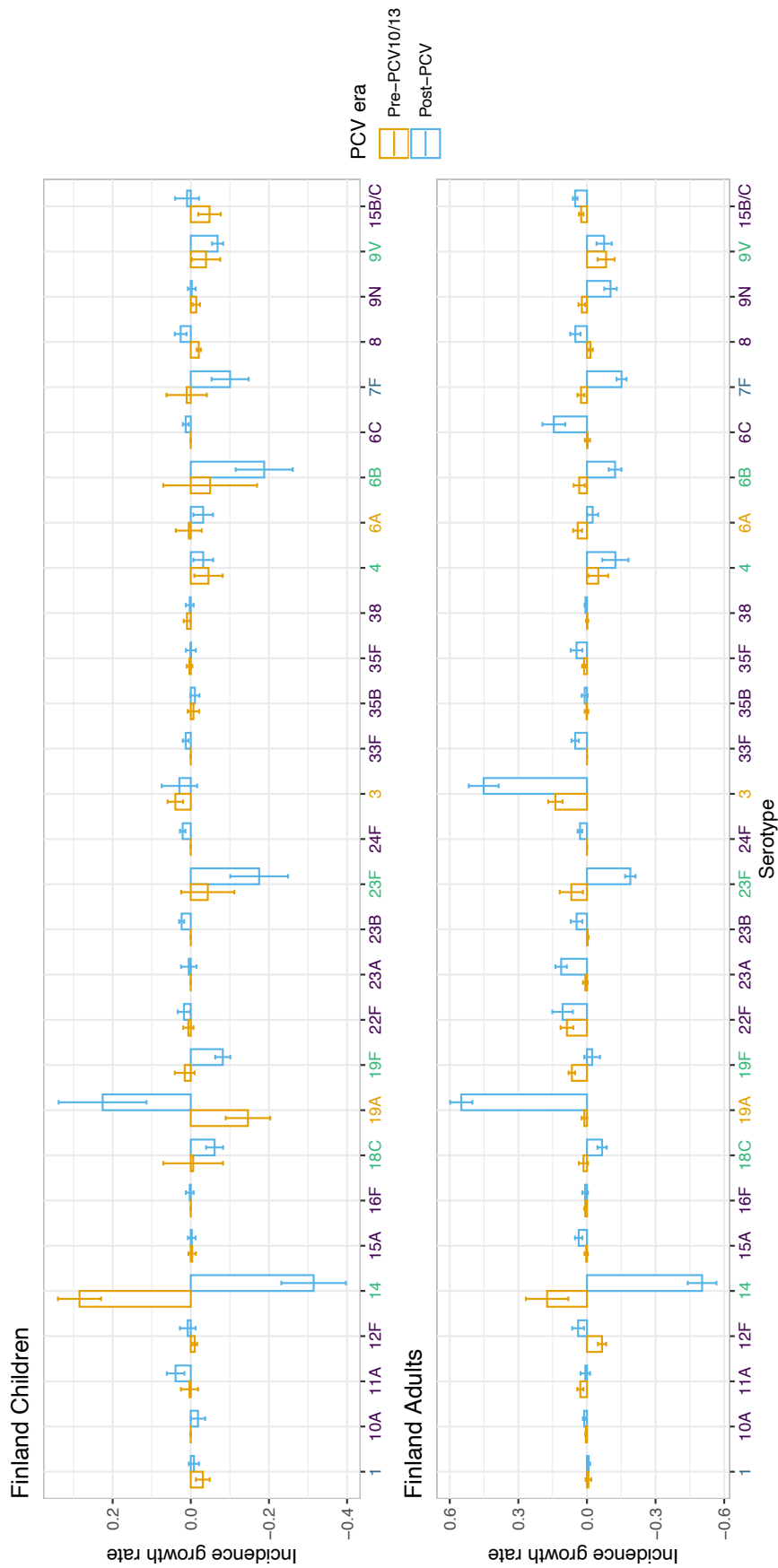
Figure 3.24: Rate of incidence growth in Finland IPD-causing serotypes pre- and post-vaccination. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years). X-axis labels in green: VT7 (PCV7 serotypes); x-axis labels in blue: VT10 (1, 5, 7F); x-axis labels in yellow (3, 6A, 19A); x-axis labels in purple: NVT (non-vaccine types)
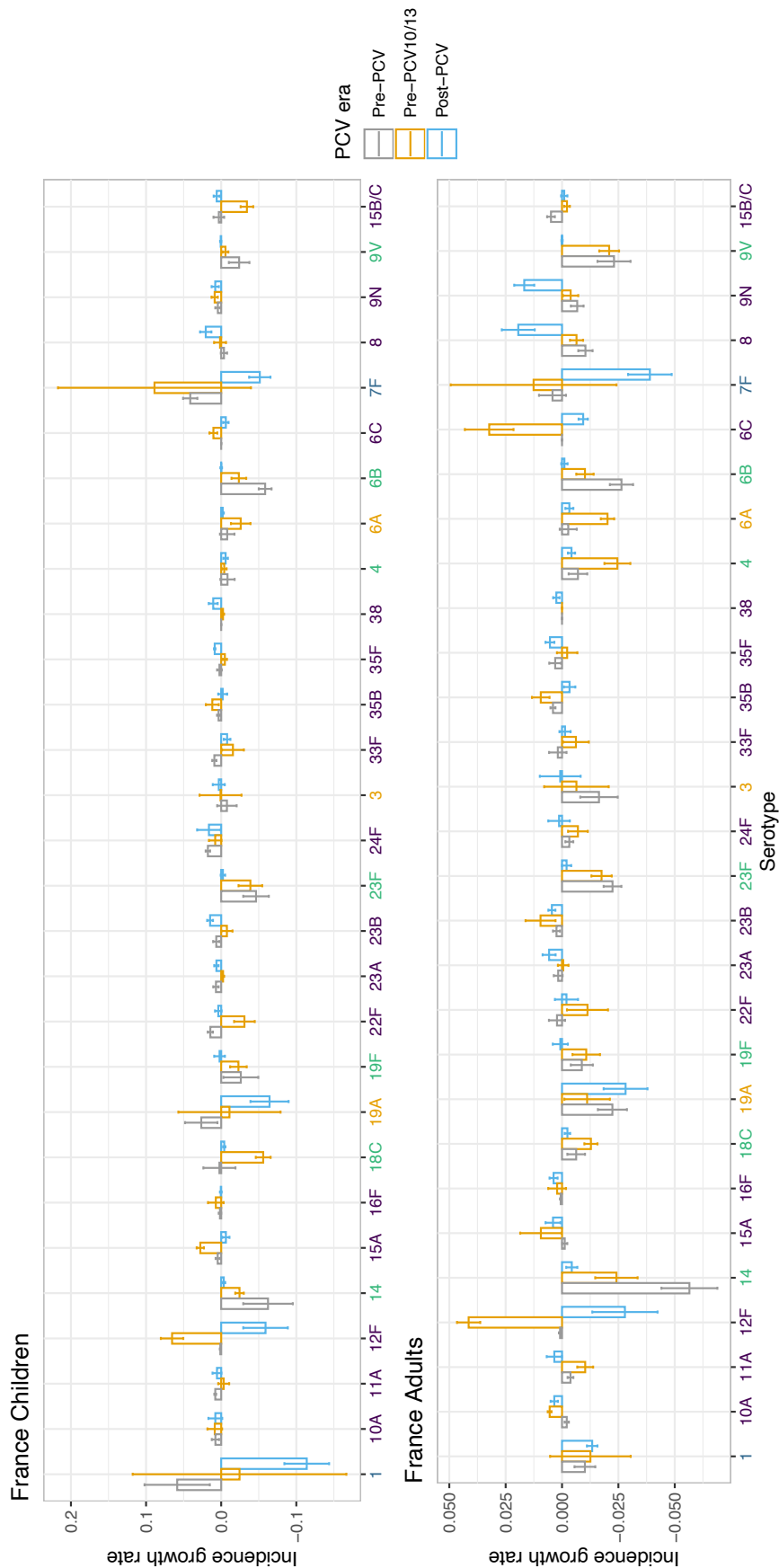
Figure 3.25: Rate of incidence growth in France IPD-causing serotypes pre- and post-vaccination. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years). X-axis labels in green: VT7 (PCV7 serotypes); x-axis labels in blue: VT10 (1, 5, 7F); x-axis labels in yellow (3, 6A, 19A); x-axis labels in purple: NVT (non-vaccine types)
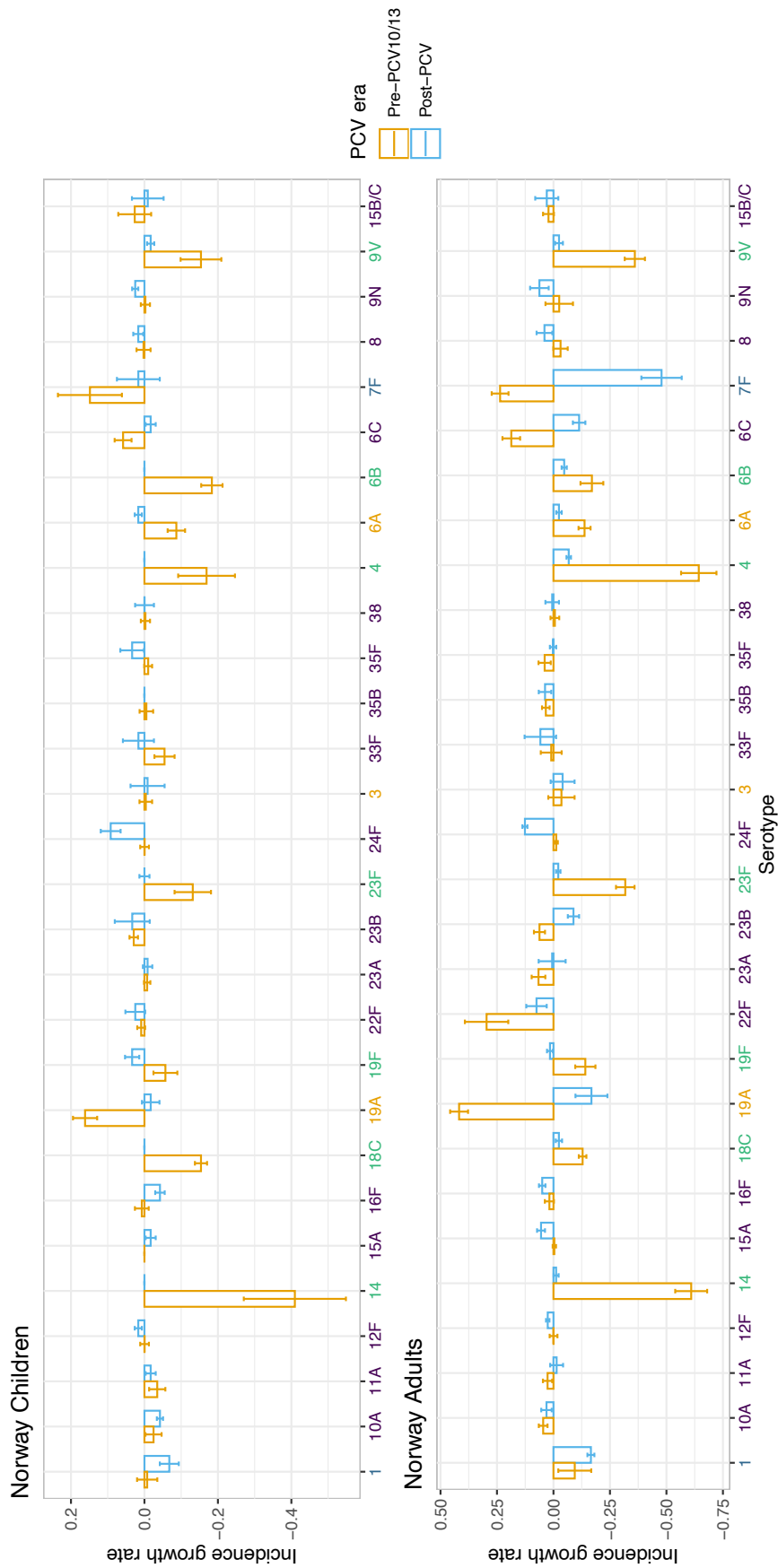
Figure 3.26: Rate of incidence growth in Norway IPD-causing serotypes pre- and post-vaccination. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years). X-axis labels in green: VT7 (PCV7 serotypes); x-axis labels in blue: VT10 (1, 5, 7F); x-axis labels in yellow (3, 6A, 19A); x-axis labels in purple: NVT (non-vaccine types)
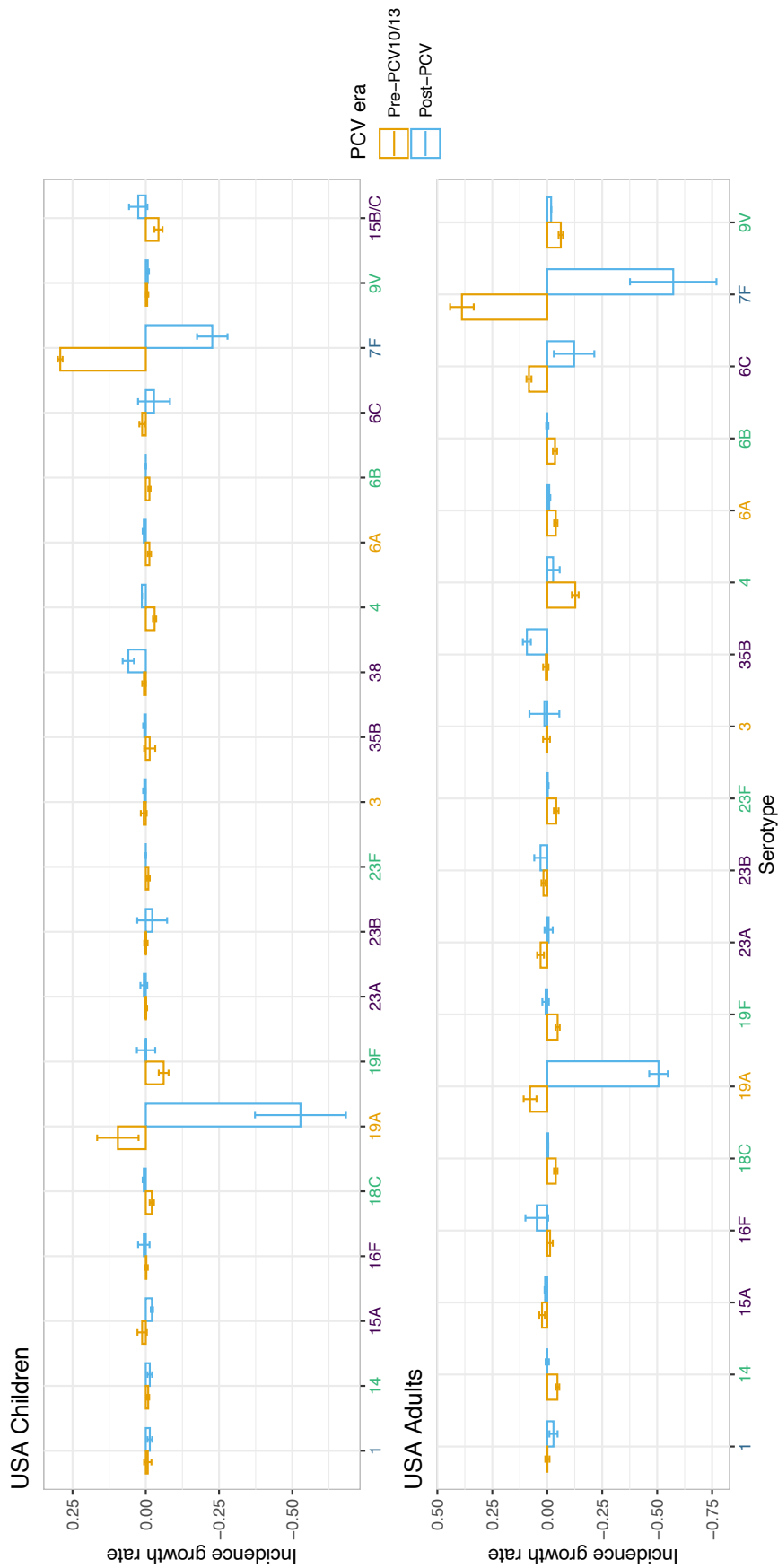
Figure 3.27: Rate of incidence growth in USA IPD-causing serotypes pre- and post-vaccination. Incidence growth rates indicate the coefficient of linear regression, i.e. the rate at which the incidence (measured as disease cases per 100,000 people per year) changes over time (measured in years). X-axis labels in green: VT7 (PCV7 serotypes); x-axis labels in blue: VT10 (1, 5, 7F); x-axis labels in yellow (3, 6A, 19A); x-axis labels in purple: NVT (non-vaccine types)

3.3 Discussion

I have analysed serotype trends pre- and post-vaccination with PCVs in selected high-income countries in Europe, North America and Australia. These have available good-quality surveillance data that enables analysis of the serotypes causing the most IPD between settings and age groups, as well as allowing identification of serotypes that are growing in incidence post-vaccination. Following the introduction of expanded-valency vaccines, certain serotypes were common in all the settings included, resulting from the persistence of VT 19A and 3 and the emergence of NVTs 12F and 23B. However, in line with other studies [326], the serotype landscape between settings has diverged post-vaccination as a consequence of serotype replacement. As such, there are no specific dominating serotypes common to all countries, as was the case pre-vaccination with PCV7 or PCV10/13. I set out to compare the differences between countries relative to which vaccines were used. These comparisons may have been confounded by reporting inconsistencies between the national surveillance datasets. The USA data only listed the top 5 most common NVTs, and consequently the results of serotype-specific IPD as a percentage of total disease are overestimated (Figures 3.11, 3.12, 3.13, 3.14), results of NVT ORs are underestimated and VT ORs overestimated (Figure 3.2) because of this omission of NVTs. Because it is the less abundant NVTs that are not reported, this should not alter the results significantly. Unfortunately, Finland was the only included country to introduce PCV10, and so systematic comparisons of PCV10 and PCV13 effects were not possible. Additionally, France and Australia were the only countries with pre-PCV7 data readily available, and there may have been other major serotypes in the pre-PCV7 era of other countries that were not captured. Although the aim was to analyse the serotype landscape in highly developed settings, it is obviously important to compare the results reported in this analysis with lower income settings, where the IPD burden is highest. For instance, serotype 5 (VT10) is an important cause of IPD in Africa but did not appear in any of the countries' rankings lists before or after vaccination. This emphasizes both the need for high-quality surveillance data from lower income countries, and the challenges for designing vaccines for worldwide use.

I also aimed to assess patterns of relative abundance in IPD pre- and post-PCV. The OR results
mirror serotype replacement observations from previous reviews [167, 168, 172, 173, 198, 327–
329], indicating the proportion of IPD-causing NVT post-vaccination increased in all countries.
The data suggest that herd immunity is also being established following the introduction
of the higher-valency PCVs, although this process may not yet be complete. In Australia
post-PCV13, VT10 were associated with a significantly greater proportion of disease relative
to other serotype categories in adults aged 18–49 years, whereas they were associated with a
significantly smaller proportion of disease relative to other serotype categories in children
less than five. On the other hand, VT13 were associated with a smaller proportion of disease
in both children less than 5 and all adults post-PCV13, suggesting herd immunity is being
established. In Finland post-PCV10, the proportion of IPD caused by VT10 did not decrease
significantly in children (<5 years), but it was reduced in adults, which is likely a sign of
herd immunity. The diversity analysis showed that while adult IPD was always caused
by a relatively wide variety of serotypes, the serotypes causing IPD in children became
increasingly diverse after vaccination. The results are consistent with the rank frequency
distributions where pre-PCV, few serotypes (i.e. VT7) were causing a majority of disease in
children: in particular serotype 14 pre-PCV in Australia and Finland, and 19A pre-PCV13
in Australia, France, and the US. However, this was not the case in adults pre-PCV, nor in
children in Norway. This absence of dominating serotypes is now the situation for both
children and adults post-PCV10/13, as children's and adults' SDI became similar in most
countries. However, Norway remains an exception, as the top-ranked serotype (22F) caused
18% of adult IPD. As the child serotype populations become more diverse, it may become
harder to achieve substantial IPD decreases with small increases in PCV valency.

Notably, serotype 3 was also consistently high in some countries after vaccination with PCV13.
This may be attributable to a lower immunogenicity in the vaccine against this serotype, or
this serotype's propensity to be carried in older individuals, who are unvaccinated [330].
Targeting this serotype more effectively is important given its ubiquity across the countries in
this study. It is not possible to quantitatively compare the benefit of targeting this serotype

more effectively over expanding valency without accounting for any associated replacement, which would be expected to be small given the low colonisation rates [203]. However, the unusual properties of serotype 3's age distribution could complicate the prediction of replacement effects in older adults. This begs the question of whether different valencies in vaccines for children and adults may be a valid approach reflecting the broad serotype distribution in geriatric age groups. While the policy around vaccinating adults is still a contentious issue because of the disease-blocking nature of the vaccine's impact rather than herd immunity-inducing nature, there is benefit in targeting adult IPD serotypes. Adult-specific vaccines would target disease and therefore circumvent the issue of replacement in carriage and avoid the potential replacement with a serotype of a higher invasiveness [331]. They may entail a huge financial commitment both from manufacturers as well as governments who may not view implementation as cost-effective depending on the cost of vaccination. Regardless of the multi-faceted considerations (public health benefit, economic, implementation, manufacturing, regulations and clinical trials), the potential public health benefit of reducing disease without causing serotype replacement is most crucial. The other option would be to vaccinate children against adult IPD serotypes in order for adult and geriatric populations to benefit from the indirect effects of childhood vaccination. However, removing adult IPD serotypes may create a space in the carriage niche of children that could be filled by more invasive (i.e. disease-causing) serotypes. This may be the preferable option for vaccine manufacturers who need only continue expanding valency in their current vaccines, although it is a strategy that presents many of the long-term challenges with replacement which we are faced with now after having twice expanded PCV valency. Unfortunately, even if serotype 3 was removed from circulation, there would still be a diverse range of serotypes to target in each country.

Comparing serotype growth and decline in age groups, one of the trends that emerged was that certain serotypes were common in one age group but rare and growing in incidence in the other age group. Examples of this are seen in Finland (serotype 11A, common in adults, is increasing in children), France (serotypes 8 and 9N, common in adults, are increasing in

children; serotypes 10A and 23B, common in children, are increasing in adults), and Norway (serotype 24F, common in children, is increasing in adults; serotypes 8 and 9N, common in adults, are increasing in children). Although it is unclear from the disease data whether adults may be a source of transmission in the community, the results reflect the known transmission between the two age groups, but may also suggest that serotype replacement is not yet complete. However, this is unclear without modelling the transmission and invasiveness of serotypes across age groups, which together would clarify whether these low-frequency but growing disease-causing serotypes are a result of replacement or age-specific invasiveness. Unfortunately, data on age-specific invasiveness are challenging to obtain at present as the carriage patterns in adults are not well understood, although additional data on adult IPD relative to infant carriage would be helpful in avoiding some of the adverse effects of paediatric immunisation programs on adult invasive disease.

The analysis of temporal changes in serotype abundance highlighted certain NVTs that appeared to be increasing significantly in multiple countries, such as serotypes 8, 9N, 15A, 23B in both young and older age groups. It is notable that no serotype has yet dominated IPD post-PCV in all countries as did serotype 14 pre-PCV and serotype 19A post-PCV7/10. There were some cases of NVTs decreasing without direct vaccination pressure, such as serotype 15A in adults in France, Norway and the USA. This may be a result of poorly understood ecological or serotype competition effects that were not explored in this analysis.

Finally, I aimed to assess the expansion of replacing NVT serotypes compared to those targeted by the vaccine and identify potential candidates for serotypes to be included in new PCVs. Taken altogether, the odds ratio, SDI and rank frequency results demonstrate that the replacing NVT serotypes are a diverse set, particularly in children where the loss of a dominating serotype in most locations post-vaccination is reflected in the increased diversity index. While vaccines tend to be implemented globally, the pooled SDI analysis for Australia, France and Norway indicates that there may be diminishing returns for adding more serotypes to new vaccines, as they will likely be proportionately less effective at reducing global IPD. This less skewed serotype distribution across IPD, described in our

results and those reported elsewhere [332, 333], is exacerbated by the post-PCV differences between countries and age groups. This makes it harder to tackle the remaining burden of disease with additional valency compared to previous PCVs. To complicate matters further, SDI, and therefore population structure, appears to be more similar between equivalent age groups of different countries. In contrast, the common serotypes appear to be more similar within countries across age groups. Hence each additional serotype added to PCVs can contribute only a relatively small further reduction in infant IPD, even in the absence of serotype replacement, and may only make such a contribution in a limited number of settings. This raises the possibility of requiring different vaccine serotype formulations for different locations, in line with ideas presented in a previous study [331], as well as distinct vaccine serotype formulations targeting adult IPD to prevent the replacement that is expected when expanding PCVs targeting infant carriage [331], as previously noted. These factors make future universal PCV development and manufacturing challenging. Although formulations could in principal vary between locations (despite this likely generating increased costs for vaccine manufacturers), optimizing PCVs for both resource-rich and resource-poor countries should be a high priority. Furthermore, despite having explored the trends of serotype replacement in disease, serotype epidemiology and temporal dynamics is greatly influenced by carriage, which is more difficult and costly to record. Understanding the rate at which certain serotypes cause disease given carriage may help in future vaccine design and in models predicting future serotype disease trends in various settings. For this, it is vital for public health officials to continue recording surveillance data post-vaccination and making it publicly available to understand both whether serotype replacement is complete post-PCV and comparisons in trends between countries.

3.4 Methods

High-income countries (as defined by the World Bank [334]) in Europe and North America with populations greater than 5 million people were chosen if they had a high PCV coverage (> 80%) and serotype-specific IPD national surveillance data publicly available from public

health agencies or published articles both before and after PCV vaccination was introduced
(Figure 3.28) for all age groups, regardless of valency, vaccination schedule or number of
surveillance years. Australia was also included as it has similar demographics to European and North American countries and to expand the analysis with its additional dataset.
Serotypes that were not typed or unknown were removed from the analyses.
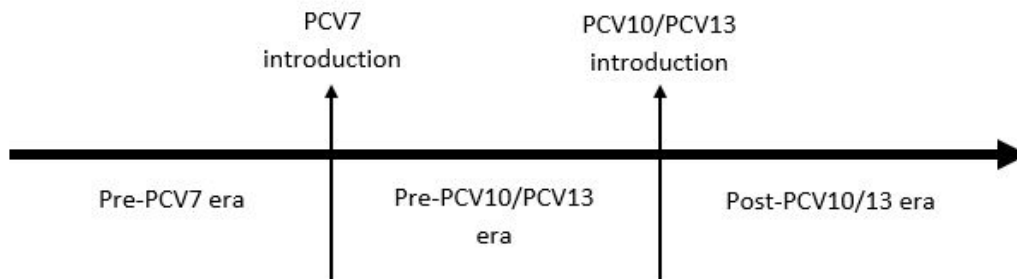
Figure 3.28: Timeline of PCV implementation eras.

3.4.1 Statistical analysis

Statistical analyses were undertaken using R. An age group-specific odds ratio (OR) and
its 95% confidence interval (CI) were calculated to assess the correlation between a vaccine
employed and an increase or decrease in the proportion of disease caused by a serotype
category in each age group. This was done by comparing the difference in disease caused
by one serotype category after vaccine introduction with the disease caused by all other
serotype categories and PCV eras. Serotype categories were PCV7 serotypes (VT7: 4, 6B,
9V, 14, 18C, 19F, and 23F), PCV10 serotypes (VT10, additional serotypes in PCV10: 1, 5, 7F),
PCV13 serotypes (VT13, additional PCV13 serotypes: 3, 6A, 19A), and non-vaccine serotypes
NVT. Suppose, for example, this table illustrates the number of IPD cases in children caused
by each serotype category in each vaccination era:

	Pre-PCV7	Pre-PCV10/13	Post-PCV
VT7	a	b	c
VT10	d	e	f
VT13	g	h	i
NVT	j	k	l

The correlation between PCV7 and the change in proportion of disease caused by VT7 would
be comparing the following:

	Pre-PCV7	Pre-PCV10/13	Post-PCV
VT7	a	b	c
VT10	d	e	f
VT13	g	h	i
NVT	j	k	l

such that the OR would be estimated as

$$OR = \frac{a * (e + f + h + i + k + l)}{(d + g + j) * (b + c)} \tag{3.1}$$

The OR demonstrates how the proportion of VT7, VT10, VT13 or NVT-caused disease cases
changed before and after vaccination in each age group, and shows the trends in relative
incidence for different serotype categories, avoiding fluctuations in absolute prevalence
that may result from non-vaccine factors. A similar analysis was previously undertaken to
show how the carriage of serotype categories changes post-vaccination [335]. My aim was
to answer whether vaccines affected the serotype categories in the same way in each age
category and country. Since countries have implemented different vaccines, the odds ratio
estimated the proportion of these serotype categories after vaccination with whichever PCVs
were implemented. The OR also circumvents the issue of calculating age-specific incidences
when surveillance populations are aggregated from different sources (such as the United
States' Active Bacterial Core surveillance), which can be complicated by non-vaccine effects
on overall IPD incidence. Since vaccine design aims to target serotypes causing the most
disease, this analysis elucidates which serotype category to target.

Simpson's Diversity Index (SDI) [336] was calculated to assess the diversity of the serotypes
causing IPD each year for each country and age group, with 95% confidence intervals [336].
SDI is an indication of richness and evenness of a population, considering the relative abun-
dance of each species, and ranges from 0 (no diversity) to 1 (infinite diversity). Populations

with few dominating serotypes are easier to vaccinate and are less diverse than those with a number of serotypes with similar abundance. In this case, estimating SDI in a population before and after vaccination can help elucidate whether vaccination has changed the distribution of the serotypes causing diseases. As such, I also estimated SDI for pooled countries four years post-vaccination, to understand whether vaccination was associated with divergence between countries' IPD serotype compositions, an indication of whether vaccines targeting additional serotypes were favourable in multiple countries.

To identify the individual serotypes underlying the OR and SDI results in each country, a rank frequency distribution for each vaccination era and country was plotted, with serotypes ranked according to the number of IPD cases caused, as well as a cumulative frequency curve, indicating a serotype's contribution to the total disease burden. Rank frequency distributions can describe population structures of the serotypes causing the most disease, particularly if they belong to a specific vaccine category (VT7 in PCV7, VT10 in PCV10, VT13 in PCV13, or NVT). Aggregation of disease cases per vaccination era allowed me to understand which serotype caused the most disease during different periods, making it insensitive to changes between individual years. These plots highlighted dominating serotypes, which I defined as the highest-ranked serotype(s) causing at least twice the burden of disease as the next most highly-ranked serotype. While the categories provide an overview of the vaccines' impacts, each serotype within the categories may behave differently as well. For this reason, visualising each individual serotype provided a more granular outlook that may elucidate the detailed patterns of serotype replacement.

To that end, the top ten IPD-causing serotypes in each country in each period were examined at each time point to compare between countries and age groups. The proportional contribution of these major serotypes were estimated as the disease cases caused by the individual serotype over total disease cases in order to compare the disease contribution between countries. This was used and plotted instead of incidence rates, as overall incidence rates can differ between countries, potentially making the disease contributions based on incidence misleading. For example, country A and country B with the same number of people

in the population and the same number of cases of serotype X would have the same incidence for serotype X, but that serotype may make up a different proportion of total cases in the population if overall incidence rates differ. In other words, the incidence does not give a measure of the contribution to the disease burden relative to the other serotypes preset. This is important for the serotype-targeting feature of PCVs, which aim to target the serotypes causing the highest proportion of disease.

Additionally, a linear regression model was employed to obtain an estimate of the crude rate at which serotype incidence was growing or declining in each vaccination period for each country in order to understand which serotypes were growing or declining, and p-values were adjusted for multiple comparisons using the Benjamini and Hochberg method [337]. Serotypes that were both ranked in the top ten and also had a high incidence growth rate post-PCV10/13 were considered important for possible future inclusion in increased valency vaccines.

3.5   Aim in this thesis

I have shown that increasing vaccine valency may be increasingly difficult for vaccine manufacturers given the observed diversity of the prevalent serotypes between countries. The analyses show that NVT serotypes 8, 9N, 15A and 23B are increasing in Australia, North America and certain European countries, even though their incidence trends are not consistent and there is a variety of other NVTs affecting each country and age group. Additionally, although there are common emerging serotypes between countries and within age groups, there has not been a dominating serotype post-PCV as there was pre-PCV (serotype 14) and pre-PCV10/13 (serotype 19A). While this chapter focused on disease incidence, transmission of pneumococcus is through carriage and therefore pneumococcal dynamic transmission models require a translation between these two important pathological events. As such, the next chapter will focus on estimating the rate at which serotypes cause disease given carriage, or the invasiveness, which is vital for future vaccine design and the model to predict future

serotype disease trends.

Chapter 4

Estimating serotype invasiveness using a Bayesian framework for translating carriage into disease

This chapter estimates and compares serotype invasiveness across datasets. Invasiveness is an estimate of a serotype's attack rate or pathogenesis, i.e. the ability of a serotype to progress from carriage to disease. In this chapter, I estimate a local invasiveness of serotypes in each dataset and a global serotype-specific invasiveness across datasets. Since transmission is through carriage, invasiveness rates allow translation of carriage prevalence estimates from dynamic transmission models to invasive pneumococcal disease incidences which are observed in populations.

Dissemination

A modified version of this chapter has been presented at Epidemics[7] - 7th International Conference on Infectious Disease Dynamics in Charleston, North Carolina, USA, 3 - 6 December 2019 as:

Løchen, A., Truscott, J.E., & Croucher, N.J. *A Bayesian approach for predicting pneumococcal serotype-specific invasiveness in children and adults in global settings.*

A modified version of this chapter is being prepared for publication in PLOS Computational
Biology as:

Løchen, A., Truscott, J.E., & Croucher, N.J. *Analysing pneumococcal invasiveness using Bayesian
models of pathogen progression rates.*

4.1   Introduction

Given the importance of the pneumococcal capsule to virulence [25, 312], and the post-PCV
changes in IPD burden (Chapter 3), it has long been assumed that capsule is an important de-
terminant of pneumococcal invasiveness. Differences in invasiveness were directly observed
in mouse experiments [338, 339] but the biological process is difficult to observe in humans.
As such, one can take an epidemiological approach to estimating the rate of this process
in humans. Pneumococcal invasiveness is a special case of the progression rate, defined as
the rate at which pneumococci move from carriage to disease (including NIPD) [340]. It
describes the ability of a serotype to progress from carriage to invasive disease, essentially
a measure of a serotype's IPD attack rate, or pathogenicity. Consistent with the hypothesis
that invasiveness is driven by the capsule, surveys of NP carriage and IPD have identified
substantial disparities in serotype frequencies between carriage and disease. For example,
'epidemic' serotypes (such as 1 and 5) are rarely found in carriage, but are sometimes found
to cause IPD outbreaks [68, 69], and are therefore thought to be highly invasive [341]. Hence
the invasiveness of pneumococcal serotypes has frequently been quantified as the ratio of its
incidence in disease to its prevalence in a matched carriage survey (termed case-to-carrier
ratio or incidence risk ratio).

Substantial changes in pneumococcal epidemiology caused by PCV introduction have ren-
dered existing methods to estimate serotype invasiveness ineffective, and have made data
synthesis challenging due to vaccine effects. Many of the 100 serotypes emerged post-PCV af-
ter having been previously rare, and vary between geographic regions (Chapter 3). Estimating
serotype invasiveness requires meta-analyses of international collections and synthesis of all

available data. This maximises the sample size for each of the plethora of serotypes found in many locations. It also minimises the serotype's association with a single genetic background, as the strain composition of different locations varies considerably [184], emphasising the need for data across locations. This is of particular importance for those serotypes that are currently rare but could emerge as important causes of disease following vaccine-associated serotype replacement.

Two methods have been used frequently to estimate the potential for serotypes to cause invasive disease. The first uses an odds, or risk, ratio relative to all other serotypes in the population [65, 342]. A variation of this has also been used in which the geometric mean of the odds ratio is taken [169]. However, when comparing multiple studies, this is confounded by variation in the mixture of other serotypes, which may not necessarily be comparable between datasets [10]. This is particularly true when comparing pre- and post-PCV studies [343], as also seen in the previous chapter (Chapter 3). The second method avoids this problem by estimating invasiveness relative to a standard serotype common in both carriage and IPD [66, 344]. However, such serotypes are likely to be targeted by PCVs as they are highly represented in carriage and disease samples across countries. Correspondingly, the original standard, serotype 14, has been eliminated by PCV7 in many settings [65]; PCV13 has similarly removed the replacement standard serotype used in post-PCV7 studies, 19A, from many locations [344]. A third method employed by one study estimated case-carrier ratios from direct calculations of IPD incidence and carriage prevalence estimates, though the authors assumed the uncertainty on the case-carrier ratio derived solely from the carriage prevalence and only estimated the ratios for one population [345]. A fourth method employing a Bayesian framework, has also been used to estimate invasiveness across PCV periods, although this was also done for only one population [223].

Yet a causal link between pneumococcal serotype and invasiveness has yet to be comprehensively established. In *H. influenzae*, changes to an isolate's serotype altered its virulence in an animal model of disease in such a manner that reflected the epidemiology of human disease [346]. While some equivalent experiments in *S. pneumoniae* have replicated observations from

human IPD [347, 348], others have found changes in an isolate's serotype did not change its
invasiveness in an animal model [349, 350], suggesting serotype-independent factors may
contribute to this phenotype. This is consistent with the Global Pneumococcal Sequencing
project's conclusion that there is variation in the invasiveness of strains of the same serotype
[351], though this is somewhat controversial [65, 66]. Previous work using genotyping data
had suggested invasiveness varied with serotype within a strain [352]. This is also consistent
with the epidemiological data, as there appear to be consistent differences between some
serotypes around the world and serotype is strongly correlated with genetic background in *S.
pneumoniae* [65, 183], particularly for highly-clonal epidemic serotypes, such as serotype 1
[351, 353]. If a serotype is clonal, many loci might contribute to invasiveness, but the serotype
would still strongly correlate with invasiveness as a consequence of linkage disequilibrium.
This makes it challenging to disentangle the contributions of different loci. This has been
attempted by testing for association of genetic loci with disease isolates, relative to those
from carriage, using comparative genomic hybridisation [354] and whole genome sequencing
[355], both of which have identified links between non-serotype-encoding loci and invasive-
ness. In combination, these observations are consistent with multiple loci contributing to
invasiveness, even if serotype is a major contributor to the overall phenotype.

I propose a novel methodology using a Bayesian framework to quantify invasiveness as a
serotype's hazard of progressing from carriage to disease. The intention was to compare
serotype compositions across multiple countries and avoid a measure that is relative either to
the rest of the pneumococcal population (i.e. the mix of serotypes in a given location), or to
one standard serotype. This method also enables a formal test of the hypothesis that serotype
alone determines pneumococcal invasiveness, through fitting models to multiple geographic
locations and host demographics. Both the estimates of serotypes' invasiveness, and the
extent to which these values exhibit heterogeneity across strains, countries and demographics,
are critical data for pneumococcal vaccine design, as PCV effectiveness relies on being able to
eliminate high-invasiveness strains [331, 356]. The data and methods used in this study are
made freely available to the public [357], to enable consistent collation of data from ongoing

surveillance of potentially problematic serotypes as vaccine-associated serotype replacement continues.

4.2 Methods

The purpose of this chapter was to develop a Bayesian framework for estimating serotype invasiveness within each published dataset reporting serotype-specific carriage and disease episodes (local invasiveness) and across these datasets (global invasiveness). A summary of the data and how it relates to model parameters can be found in Figure 4.1.

Data
- Number of carriers of each serotype i in dataset j (c_{ij})
- Number of disease cases of each serotype i in dataset j (d_{ij})
- Total number of nasopharyngeal swabs in dataset j (η_j)
- Total disease population in dataset j (N_j)
- Time period of study in dataset j (t_j)

# carriers	Total carrier swabs	# IPD cases	Total population, N	Length of study, t
⋮	⋮	⋮	⋮	⋮
n datasets	n datasets	n datasets	n datasets	n datasets

Bayes' theorem rearranged:
$$P(d, v, \rho) \propto P(d|v, \rho)\, P(v|\rho)\, P(\rho)$$
where d = # disease cases, v = invasiveness (disease cases per 100,000 carriers-year), ρ = carriage prevalence.

Priors

$P(v|\rho)$: prior on invasiveness, independent of carriage prevalence, flat/uninformative

$P(\rho)$: prior on carriage prevalence, beta distribution given # of carriers follows binomial distribution

$$c_{ij} \sim \text{Binom}(\rho_{ij}, \eta_j)$$
$$P(\rho_{ij}) = \text{Beta}(\rho_{ij}, \alpha = c_{ij} + 1, \ \beta = \eta_j - c_{ij} + 1)$$

Likelihood

$P(d|v, \rho)$: likelihood of disease data, Poisson distribution with disease rate $\lambda = v\rho Nt$ which follows a gamma distribution

$$\lambda \sim \text{Gamma}(\alpha = d_{ij} + 1, \beta = 1)$$

Local invasiveness posterior distribution
$$v_{ij} = \frac{\lambda_{ij}}{\rho_{ij} N_j t_j}$$
Estimated from deviates of λ_{ij} from gamma distribution and deviates of ρ_{ij} from beta distribution

Global invasiveness (logarithmic) posterior distribution
$$\ln P(v_i|\{d_{ij}\}) \cong \sum_{j=1}^{n} \ln\left(\frac{1}{M}\sum_{k=1}^{M} P(d_{ij}|\lambda_{ij}(\rho_{ij}^k))\right)$$
Global invasiveness approximated using M deviates from the beta-distributed ρ_{ij}.

Figure 4.1: Summary of how data from published datasets links to model parameters for the estimation of local and global invasiveness.

4.2.1 Dataset selection

From a previously published meta-analysis [331], datasets reporting serotype-specific numbers of carriers and disease cases of children and adults were gathered. Briefly, a PubMed search on 5 October 2017 by the study authors using the terms "(case[All Fields] OR disease[All Fields] OR episode[All Fields] OR patient[All Fields]) AND (carriage[All Fields] OR

carrier[All Fields] OR nasopharyngeal[All Fields]) AND (invasiveness[All Fields] OR "attack

rate"[All Fields] OR "type distribution"[All Fields] OR "serotype distribution"[All Fields]

OR "serogroup distribution"[All Fields] OR "invasive capacity"[All Fields] OR "invasiveness

ratio"[All Fields] OR "odds ratio"[All Fields] OR "carrier ratio"[All Fields] OR ("invasive iso-

lates"[All Fields] AND "carriage isolates"[All Fields])) AND ("serogroup"[MeSH Terms] OR

"serogroup"[All Fields] OR "serotype"[All Fields]) AND ("streptococcus pneumoniae"[MeSH

Terms] OR ("streptococcus"[All Fields] AND "pneumoniae"[All Fields]) OR "streptococcus

pneumoniae"[All Fields] OR "pneumococcus"[All Fields])" retrieved all studies studying

pneumococcal serotype invasiveness with both serotype-specific distributions in carriage and

invasive disease [331].

The literature search returned 136 results, of which 32 were looked at in detail after reviewing

the papers. After contacting authors, twelve were further rejected for the following reasons:

- Study design was biased towards particular serotypes (Takala et al. [358])

- Very high ($> 75\%$) co-colonisation levels made the population structure of the carriage
  sample difficult to ascertain (Azzari et al. [359])

- Could not access the raw data for re-analysis (Valles et al. [360])

- No high-resolution serotype data reported (Principi et al. [361])

- No clear age stratification of the collection (Smart et al. [362], Galanis et al. [363])

- Very small sample sizes once stratified by age (Cullotta et al. [364])

- Could not access manuscript (Mogdasy et al. [365])

- No information on total nasopharyngeal swabs obtained (Scott et al. [366], three studies
  reported in Brueggemann et al. [65], Browall et al. [367])

- Absence of data on the total population (Levidiotou et al. [368])

- 'True' carriage prevalence difficult to ascertain due to selection of positive isolates
  (Hanage et al. [369], one study reported in Brueggemann et al. [65]): From the total
  number of positive swabs, a random selection of isolates was used for the analysis. Of
  these, isolates were further selected in Hanage et al. as not all were true positives for *S.
  pneumoniae*, and further selected again for the relevant time period.

- Serotype reporting in carriage surveillance did not differentiate between vaccinated
  and unvaccinated children (Mameli et al. [370], Løvlie et al. [371])

This left 21 datasets with substantial matched systematically-sampled and thoroughly serotyped asymptomatic carriage and disease samples [192, 223, 342, 352, 372–384] (Table 4.2). Within these, isolates of serotypes 15B and 15C were combined into the single 15B/C category. Samples were stratified by age and date of vaccine introduction.

4.2.2 Construction of the serotype invasiveness model

Bayes' Theorem states that the probability of A given B is

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{4.1}$$

where $P(B|A)$ is the likelihood, i.e. the probability of B being true given A, $P(A)$ is the prior (i.e. the known) probability of A, and $P(B)$ is the probability of B. $P(B)$ is typically not something that is observed, and therefore is considered a nuisance parameter to be integrated out, also known as marginalisation. This process can be done by summing the probabilities of all the possibilities of B, or taking the integral.

I developed a Bayesian framework to estimate the invasiveness of each serotype from the data for each age group (children and adults). Invasiveness is defined as the rate of disease given carriage (i.e. disease cases per carrier per year). In the first model, I estimated a separate invasiveness v_{ij} for serotype i in dataset j (i.e. "local" invasiveness) such that each dataset can be treated independently. Inference of invasiveness is based on two independent sources of data; carriage prevalence of serotypes in a population, ρ_{ij}, and disease cases detected in the same population, d_{ij}. From Bayes Theorem (Equation 4.1), the joint probability of the data, the invasiveness and the carriage prevalence can be written as

$$P(d_{ij}, v_{ij}, \rho_{ij}) = P(d_{ij}|v_{ij}, \rho_{ij})\pi(v_{ij}|\rho_{ij})\pi(\rho_{ij}) \tag{4.2}$$

where functions $\pi(\cdot)$ represent the priors. The prior for the invasiveness, $\pi(v_{ij})$, is assumed to be flat and independent of the carriage prevalence and can therefore be omitted. ρ_{ij} can be estimated from population surveys. For serotype i in population j, c_{ij} is the number of positive swabs detected in a sample of size η_j, implying a carriage prevalence of ρ_{ij}. As such, c_{ij} is a result of a binomial process reflecting the background carriage prevalence:

$$c_{ij} \sim \text{Binom}(\rho_{ij}, \eta_j) \tag{4.3}$$

If the prior distribution for prevalence is a uniform distribution between 0 and 1, the posterior will be beta-distributed:

$$\pi(\rho_{ij}) = \text{Beta}(\rho_{ij}; \alpha = c_{ij} + 1, \beta = \eta_j - c_{ij} + 1) \tag{4.4}$$

Furthermore, the first term on the right in Equation 4.2 is the likelihood of the disease count data, d_{ij}, which are considered to be the result of a Poisson process in response to a force of infection for a given serotype and a given population:

$$P(d_{ij}|v_{ij}, \rho_{ij}) = \text{Poisson}(d_{ij}; \lambda(v_{ij}, \rho_{ij})) = \frac{\lambda(v_{ij}, \rho_{ij})^{d_{ij}}}{d_{ij}!} \exp(-\lambda(v_{ij}, \rho_{ij})) \tag{4.5}$$

where $\lambda(v_{ij}, \rho_{ij}) = \lambda_{ij}$ is the expected count. The disease incidence is assumed to be constant over the period of observation, therefore λ_{ij} can be expressed in the following way:

$$\lambda_{ij} = \lambda(v_{ij}, \rho_{ij}) = v_{ij}\rho_{ij}N_j\Delta t_j \tag{4.6}$$

where N_j is the human population size under surveillance in the dataset and Δt_j is the time period over which the cases were detected. The quantities N_j and Δt_j are assumed to be known with negligible error. v_{ij} is the invasiveness, which can be interpreted as the probability per unit time (or hazard) at which an individual carrying a serotype will develop the disease of interest, in this case IPD. The posterior distribution results from integrating out the nuisance parameter ρ_{ij} and is given by

$$P(v_{ij}|d_{ij}) \propto \int_{\rho_{ij}} P(d_{ij}|\lambda(v_{ij}, \rho_{ij}))P(\rho_{ij}|c_{ij}, \eta_{ij})d\rho_{ij}. \tag{4.7}$$

The simplest way to calculate the statistics of v_{ij} is through direct calculation from deviates, i.e. independent and identically distributed random variables drawn from the distributions of the other parameters. Since the disease count data are distributed according to the Poisson distribution, the expected count is distributed according to the conjugate prior distribution, the Gamma distribution:

$$\lambda_{ij} \sim \text{Gamma}(\alpha = d_{ij} + 1, \beta = 1). \tag{4.8}$$

So, ignoring the uncertainty in N_j and Δt_j, drawing independent deviates of λ_{ij} and ρ_{ij} from the Gamma distribution and the Beta distribution respectively can generate deviates of v_{ij} rearranging Equation 4.6:

$$v_{ij} = \frac{\lambda_{ij}}{\rho_{ij}N_j\Delta t_j}. \tag{4.9}$$

In the second model, I estimated an invasiveness for each serotype that draws on all datasets in which it is present, i.e. an invasiveness v_i for each serotype i in n datasets (i.e. "global" invasiveness). In this case, the different datasets cannot be separated and the posterior distribution (Equation 4.7) takes the form

$$P(v_i|\{d_{ij}\}) \propto \prod_{j=1}^{n} \int_{\rho_{ij}} P(d_{ij}|\lambda(v_i, \rho_{ij}))P(\rho_{ij}|c_{ij}, \eta_{ij})d\rho_{ij}. \tag{4.10}$$

To marginalise over the nuisance parameter ρ_{ij}, the integral can be approximated using M

independent deviates from the Beta distribution of ρ_{ij}, ρ_{ij}^{k}, where $k = 1 \ldots M$

$$\int_{\rho_{ij}} P(d_{ij}|\lambda(v_i, \rho_{ij}))P(\rho_{ij})d\rho_{ij} \approx \frac{1}{M}\sum_{k=1}^{M} P(d_{ij}|\lambda(v_i, \rho_{ij}^{k})) \tag{4.11}$$

i.e. the logarithmic posterior probability of the invasiveness of serotype i, for a set of values

of disease cases across all datasets $\{d_{ij}\}$ can be written as:

$$\ln P(v_i|\{d_{ij}\})) \cong \sum_{j=1}^{n} \ln \left(\frac{1}{M}\sum_{k=1}^{M} P(d_{ij}|\lambda_{ij}(\rho_{ij}^{k}))\right) \tag{4.12}$$

An arbitrary value of $M = 10,000$ was sufficient to produce minimal variance in the integral,

since a minimum of 1000 samples are needed to construct a posterior distribution and build a

credible interval [385, 386]. In this way, the posterior distribution of v_i can be evaluated to

any degree of accuracy. The credible intervals (CrIs) in our case were estimated from 95% of

the highest posterior density region. All analyses were undertaken using R.

The following assumptions were made in all invasiveness models:

- The population, carriage prevalence and disease hazard were constant throughout the
  years included in the studies.
- Unless otherwise stated, the disease surveillance was assumed to cover 100% of the
  population specified in the respective manuscripts.
- There were two age groups: children ($\leq$ 18 years old) and adults ($>$ 18 years old). This
  stratification reflected that of the meta-analysis datasets. Most of the cases in children $<$
  18 years can be expected to be in children $<$ 5 years.
- Adult carriage was assumed to be the same as children's, as it was more readily available.

Adult upper respiratory carriage is thought to be underestimated and therefore may
not be reliable [387]. Furthermore, vaccine serotypes are typically eliminated from adult
IPD, except in cases where adult carriage is also high [55, 379].

- Each carried isolate came from an individual person, and multiple serotypes were
counted as separate carriers. As such, carriage rates are effectively independent of
one another, and therefore could sum to more than one in principle, given sufficient
multiple carriage.

- There are no false negative swabs; i.e., detection of all serotypes is effectively perfect,
regardless of which method for serotyping was employed.

4.2.3 Model comparison

The obtained posterior distributions were qualitatively compared to those obtained by Weinberger et al. [223], who also used a Bayesian approach with only one population (Navajo and White Mountain Apache populations in the US) across vaccination periods, to understand whether their results should be used as informative priors for our posterior distributions. Briefly, their modelling framework estimated serotype-specific invasiveness for each vaccination period, age group and serotype in the same fashion as presented here, but pooled estimates across vaccination periods by assuming the period-specific estimates were drawn from a normal distribution with a mean equivalent to the 'true' invasiveness and the precision equivalent to the period-specific precision. This pooled invasiveness estimate was compared to both the local and global invasiveness estimates presented here.

Additionally, I aimed to quantitatively compare the local invasiveness model and the global invasiveness model. Given two competing models, one would want to compare the posterior distributions, i.e. the probability of the respective models, say M_1 and M_2, given the data D, $P(M_1|D)$ and $P(M_2|D)$. From Bayes Theorem (Equation 4.1), $P(M|D)$ can be estimated as:

$$P(M|D) = \frac{P(D|M) \cdot P(M)}{P(D)} \tag{4.13}$$

where $P(D)$ is the probability of the data, which would be the same for the two models, and $P(M)$ is the prior probability of the model. The two models can then be compared via their posterior odds as:

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(D|M_1)}{P(D|M_2)} \times \frac{P(M_1)}{P(M_2)} \tag{4.14}$$

The first term on the right-hand side is Bayes Factor (BF), which is the ratio of the probability of the data given the models, or in other words, the relative strength of evidence of one model over the other. It can also be interpreted as the degree to which one model updates the posterior based on their prior [388]. If the priors on the models are the same, then the posterior odds are equivalent to BF:

$$BF = \frac{P(D|M_1)}{P(D|M_2)} \tag{4.15}$$

If $BF > 1$, the data favour M_1, whereas if $BF < 1$, the data favour M_2. The farther from 1 (in either direction), the stronger the evidence for one model over the other:

Table 4.1: Bayes Factor evidence labelling. Reproduced from [3].

Bayes Factor	Label
> 100	Extreme evidence for M_1
$30 - 100$	Very strong evidence for M_1
$10 - 30$	Strong evidence for M_1
$3 - 10$	Moderate evidence for M_1
$1 - 3$	Anecdotal evidence for M_1
1	No evidence
$\frac{1}{3} - 1$	Anecdotal evidence for M_2
$\frac{1}{3} - \frac{1}{10}$	Moderate evidence for M_2
$\frac{1}{10} - \frac{1}{30}$	Strong evidence for M_2
$\frac{1}{30} - \frac{1}{100}$	Very strong evidence for M_2
$< \frac{1}{100}$	Extreme evidence for M_2

In this way, the size of BF allows one to understand which model is preferred based on the relative evidence.

To understand which model was preferable between the local and global invasiveness models, I estimated and labelled BF [388, 389] in order to compare the relative evidence of the local versus global models. Because the model priors are the same, BF is equivalent to the posterior odds:

$$BF_i = \frac{\prod_{j=1}^{n} P(v_{ij}|d_{ij})}{P(v_i|\{d_{ij}\})} \tag{4.16}$$

In this way, the marginal likelihood of each model can be approximated as the area under the posterior distribution curve, and BF can be estimated to establish which model the evidence favours.

4.2.4 Comparison with genetic data

To test whether variation in v_i might be explained by differential association with other genetic loci, the diversity of global pneumococcal sequence clusters (GPSCs) with which each serotypes was associated was calculated. The GPSC data is a global collection of isolates pre- and post-PCV [351], and is the most comprehensive resource available for determining the linkage between serotype and genetic background. The relationship between the number of strains and the global invasiveness estimates was quantified, using the Pearson correlation coefficient.

4.3 Results

4.3.1 Heterogeneity in datasets

The literature search resulted in 20 pairs of infant carriage and IPD samples (nine of which were post-PCV introduction), and five pairs of infant carriage and primarily adult IPD samples (one of which was post-PCV introduction) (Table 4.2). All the studies described the children's age group except for one which was adult age group only, and only five studies

described both children and adult age groups. From the twenty-one datasets gathered from

the meta-analysis, just under two thirds were from the pre-vaccination era; seven were from

the post-PCV7 era, and only two were post-PCV10/13 (Table 4.3). The number of swabs

taken in the studies ranged between 197 and 6541. Child datasets described between 7 and

47 serotypes in circulation, whereas adult datasets described between 12 and 44 (Table 4.4,

Table 4.5). VT serotypes were among those most prevalent in carriage pre-PCV, although there

is no obvious trend among datasets or serotypes (Figure 4.2, Figure 4.3), and the variance

was large between datasets.

Table 4.2: Summary of datasets included in meta-analysis

Population	Vaccination period	Time frame	Carriage isolates source	Infant disease source	Adult disease source	Number of serotypes	References
Alabama	Pre-PCV	July 1975 - December 1978	Unvaccinated sick and healthy children < 18 years old from Alabama	Unvaccinated children with IPD <18 years old from Alabama	Unvaccinated adults hospitalised with pneumonia or IPD in Alabama	Children: 11 Adults: 12	Gray et al. 1979 [373]
Goroka	Pre-PCV	1981 - 1987	Unvaccinated children attending clinics near Goroka Town	Unvaccinated children with IPD in Goroka hospital	-	Children: 21	Smith et al. 1993 [342]
Ontario	Pre-PCV	1995	Unvaccinated healthy children primarily < 4 years old in Toronto	Unvaccinated children < 18 years old with IPD in Toronto and surrounding area	-	Children: 8	Kellner et al. 1998 [375]
Atlanta	Pre-PCV	January 1995 - December 1995	Unvaccinated children with recent URI in Atlanta <5 years old	Unvaccinated children < 5 years old with IPD in Atlanta	-	Children: 10	Sharma et al. 2013 [192]

Population	Vaccination period	Time frame	Carriage isolates source	Infant disease source	Adult disease source	Number of serotypes	References
Czech	Pre-PCV	1996 - 2005	Unvaccinated healthy children 3 - 5 years old across the Czech Republic	Unvaccinated children < 6 years old with IPD across the Czech Republic	-	Children: 27	Zemlickova et al. 2010 [383]
England & Wales	Pre-PCV	July 1996 - June 2006	Unvaccinated healthy children <5 years old in Hertfordshire	Unvaccinated children <5 years old across England & Wales	Predominantly unvaccinated individuals >4 years old across England & Wales	Children: 28 Adults: 31	Trotter et al. 2010 [380]
Stockholm	Pre-PCV	1997	Unvaccinated healthy children <7 years old in Stockholm	-	Unvaccinated adults (and a small number of children) hospitalised with IPD in Stockholm	Adults: 30	Sandgren et al. 2004 [378]
Portugal	Pre-PCV	January 2001 - December 2003	Unvaccinated healthy children <7 years old in Lisbon and Oeiras	Unvaccinated children (< 18 years) across Portugal	Unvaccinated adults (> 18 years) across Portugal	Children: 35 Adults: 44	Sá-Leao et al. 2011 [352]
Netherlands	Pre-PCV	June 2004 - May 2006	Unvaccinated healthy children <2 years old in Noord-Holland, Zuid-Holland and Utrecht	Children < 5 years with IPD across Netherlands	-	Children: 28	Vissers et al. 2018 [384]

Population	Vaccination period	Time frame	Carriage isolates source	Infant disease source	Adult disease source	Number of serotypes	References
Bogota	Pre-PCV	May 2005 - November 2006	Healthy unvaccinated children <18 months old in Bogota	IPD in children <2 years old in Bogota	-	Children: 37	Parra et al. 2013 [376]
Caracas	Pre-PCV	December 2006 - January 2008	Unvaccinated healthy children <6 years old in Caracas	Unvaccinated children with IPD <6 years old in Caracas	-	Children: 12	Rivera-Olivero et al. 2011 [377]
Morocco	Pre-PCV	November 2010 - December 2011	Unvaccinated healthy children in Rabat	Children hospitalised with severe pneumonia in Rabat <5 years old	-	Children: 8	Jroundi et al. 2017 [374]
Massachusetts	Post-PCV7	2001 - April 2009	PCV7-vaccinated children visiting physicians <7 years old across Massachusetts	PCV7-vaccinated children <7 years with IPD old across Massachusetts	-	Children: 42	Yildirim et al. 2010 [382]

Population	Vaccination period	Time frame	Carriage isolates source	Infant disease source	Adult disease source	Number of serotypes	References
Navajo	Post-PCV7	March 2006 - March 2008	PCV7-vaccinated Navajo or White Mountain Apache native American children <9 years old	Active IPD surveillance of PCV7-vaccinated Navajo or White Mountain Apache native American children <7 years old	Active IPD surveillance of PCV7-vaccinated Navajo or White Mountain Apache native American children at least 18 years old	Children: 39 Adults: 40	Scott et al. 2012 [379]; Weinberger et al. 2016 [223]
Barcelona	Post-PCV7	2007 - 2011	PCV7-vaccinated healthy children <7 years old in Barcelona	PCV7-vaccinated children <7 years old with IPD in Barcelona	-	Children: 23	del Amo et al. 2014 [372]
France	Post-PCV7	January 2008 - December 2009	PCV7-vaccinated healthy children <2 years old across France	PCV7-vaccinated children with IPD <2 years old across France	-	Children: 47	Varon et al. 2015 [381]
Atlanta	Post-PCV7	June 2008 - May 2009	PCV7-vaccinated sick Georgia-resident children <5 years old	PCV7-vaccinated children <5 years old with IPD in Atlanta	-	Children: 17	Sharma et al. 2013 [192]

Population	Vaccination period	Time frame	Carriage isolates source	Infant disease source	Adult disease source	Number of serotypes	References
Netherlands	Post-PCV7	June 2008 - May 2012	Vaccinated healthy children <2 years old in Noord-Holland, Zuid-Holland and Utrecht	IPD in children (< 5 years) across Netherlands	-	Children: 37	Vissers et al. 2018 [384]
Bogota	Post-PCV7	June 2011 - November 2011.	PCV7-vaccinated healthy children <18 months old in Bogota	IPD in children <2 years old in Bogota	-	Children: 35	Parra et al. 2013 [376]
France	Post-PCV13	January 2012 - December 2013	PCV13-vaccinated healthy children <2 years old across France	PCV13-vaccinated children with IPD <2 years old across France	-	Children: 38	Varon et al. 2015 [381]
Netherlands	Post-PCV10	June 2012 - May 2016	Vaccinated healthy children <2 years old in Noord-Holland, Zuid-Holland and Utrecht	Children (< 5 years) with IPD across Netherlands	-	Children: 36	Vissers et al. 2018 [384]

Table 4.3: Summary of dataset populations. E&W: England and Wales.

	Dataset	No. of swabs	Study disease population of children	Disease cases children	Study disease population of adults	Disease cases adults	Time interval of study (years)
Pre-PCV	Alabama	827	19,316 [390]	114	232,373 [390]	86	3
	Atlanta	231	204,680	202	-	-	1
	Bogota	197	357,200	353	-	-	4
	Caracas	1004	146,125 [391]	36	-	-	1
	Czech	425	478,177 [392]	138	-	-	9
	E&W	3752	3,091,000	461	48,702,414	1876	1
	Goroka	2844	96,207 [393]	56	-	-	6
	Morocco	200	212,566 [394]	118	-	-	1
	Netherlands	321	250,924 [395]	100	-	-	2
	Ontario	1139	580,507 [396]	89	-	-	1
	Portugal	1170	2,071,223 [397]	90	8,284,894 [397]	378	2
	Stockholm	611	-	-	2,004,152 [398]	273	1
Post-PCV7	Atlanta	451	298,831	47	-	-	1
	Barcelona	209	228,000	159	-	-	4
	Bogota	246	357,200	91	-	-	1
	France	1212	838,866 [399]	388	-	-	6
	Massachusetts	2969	820,000	206	-	-	3
	Navajo	6541	65,048 [400]	132	201,553 [400]	514	6
	Netherlands	660	232,251 [395]	73	-	-	4
Post-PCV10/13	France	1212	842,076 [399]	181	-	-	6
	Netherlands	659	222,671 [395]	47	-	-	4

Table 4.4: Summary of serotypes included in each dataset reporting children.

Serotype	Pre-PCV Alabama	Atlanta	Bogotá	Caracas	Czech	Eng&Wales	Goroka	Morocco	Netherlands	Ontario	Portugal	Post-PCV7 Atlanta	Barcelona	Bogotá	France	Massachusetts	Navajo	Netherlands	Post-PCV10/13 France	Netherlands
4	×	×	×	×	×	×		×	×	×	×			×	×	×	×	×	×	×
6B		×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×	×	
9V	×	×	×	×	×	×			×	×	×		×	×	×	×	×	×	×	×
14	×	×	×	×	×	×	×		×	×	×		×	×	×	×	×	×		×
18C		×	×		×	×		×	×		×			×	×	×	×	×	×	×
19F		×	×	×	×	×			×	×	×	×	×	×	×	×	×	×	×	×
23F		×	×	×	×	×			×	×	×	×	×	×	×	×	×	×		×
1	×		×	×	×		×		×		×	×	×	×	×	×	×	×		×
5	×		×	×	×		×				×		×	×	×	×	×			×
7F			×	×	×	×		×	×	×	×	×	×		×	×	×	×	×	×
3	×		×	×	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×
6A		×	×	×	×	×		×	×	×	×		×	×	×	×		×	×	×
19A		×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	×	×	×
10A			×		×	×	×		×		×	×	×	×	×		×	×	×	×
10B			×												×					
10C															×		×			
10F															×		×			
11A		×	×		×	×			×	×	×	×	×	×	×	×	×	×	×	×
11B															×	×				
11F																				
12B																×				
12F			×		×						×		×	×	×	×	×		×	×
13	×		×								×		×	×	×	×	×		×	×
15A			×		×	×	×		×		×	×	×	×	×	×	×	×	×	×
15B/C			×			×	×		×		×	×			×	×	×	×	×	×
15F			×								×					×				
16F			×		×	×			×		×		×	×	×	×	×	×	×	×
17F						×			×		×		×	×	×	×	×	×	×	×
18A			×			×					×									
18B			×																	
18F			×													×				
19B														×						
19C											×									
2							×				×	×	×	×	×	×	×	×	×	×
20	×		×			×	×	×	×		×	×		×			×			×
21			×		×	×											×			
22A								×												
22F			×		×	×			×		×	×	×	×	×	×	×	×	×	×
23A	×		×			×	×		×		×	×	×	×	×	×	×	×	×	×
23B		×	×						×		×	×	×	×	×	×	×	×	×	×

	Pre-PCV											Post-PCV7							Post-PCV10/13	
	Alabama	Atlanta	Bogota	Caracas	Czech	Eng&Wales	Goroka	Morocco	Netherlands	Ontario	Portugal	Atlanta	Barcelona	Bogota	France	Massachusetts	Navajo	Netherlands	France	Netherlands
24A															X			X	X	
24B			X						X		X			X	X			X	X	X
24F			X								X			X	X	X		X	X	X
25A						X											X			
27					X									X						
28A	X						X								X					
28F	X				X	X	X				X			X	X					
29															X	X	X	X	X	X
31									X							X	X	X	X	X
33A									X		X				X	X	X	X	X	X
33F					X						X	X	X			X	X			
34			X				X								X		X		X	X
35A	X		X			X			X			X	X		X	X	X	X	X	X
35B	X		X														X		X	X
35C	X		X																	
35F			X		X	X					X			X		X		X		
36							X				X			X	X	X		X	X	X
37					X											X	X	X		
38	X				X	X	X		X			X			X	X	X			X
39							X													
40					X													X		
42							X							X				X		
43							X													
45							X								X					
46							X								X					
47							X								X					
6C			X			X					X	X	X	X	X	X		X	X	X
7C	X		X		X	X	X	X	X		X			X	X	X	X	X	X	
8														X	X		X			
9A	X														X	X			X	
9L									X		X				X					
9N			X		X	X			X		X				X	X	X	X	X	X
Total	11	10	37	12	27	28	21	8	28	8	35	17	23	35	47	42	39	37	38	36

Table 4.5: Summary of serotypes included in each dataset reporting adults.

	Pre-PCV				Post-PCV7
	Alabama	**Eng&Wales**	**Portugal**	**Stockholm**	**Navajo**
4	X	X	X	X	X
6B		X	X	X	X
9V		X	X	X	X
14	X	X	X	X	X
18C		X	X	X	X
19F		X	X	X	X
23F		X	X	X	
1	X		X	X	X
5		X			X
7F		X	X	X	X
3	X	X	X	X	X
6A		X	X	X	
19A		X	X	X	X
10A		X	X		X
10C					X
10F			X		X
11A		X	X	X	X
12B			X		
12F				X	X
13	X		X		X
15A			X		X
15B/C		X	X	X	X
16F		X	X		X
17F		X	X	X	X
18A			X		
18B		X			
2				X	
20	X	X	X	X	X
21		X	X	X	X
22A					X
22F		X	X	X	X
23A		X	X	X	X
23B			X		X
24B			X		
24F			X	X	
25F			X		
27		X			X
29	X				
31	X	X	X	X	X
33A					X
33F			X	X	X
34	X	X	X		X
35A		X		X	X
35B		X	X	X	X
35F		X	X	X	X
37	X	X	X		X
38	X	X	X	X	X
6C			X		
7C			X		X
8	X	X	X	X	X
9A			X		
9L			X		
9N		X	X	X	X
Total	12	31	44	30	40

About two thirds of the studies took place in high-income settings, with the rest taking place
in lower- or middle-income settings. Overall IPD incidence in study locations varied from
2 (Portugal) to 94 (Atlanta) cases per 100,000 people pre-vaccination in children, from 5.3
(Netherlands post-PCV10) to 33 (Navajo post-PCV7) cases per 100,000 people post-vaccination

in children, and from 2 (Portugal) to 13 (Stockholm) cases per 100,000 people pre-vaccination in adults (Figure 4.4). Portugal's low incidence of disease is likely attributed to its large study disease population (Table 4.3), which surveillance was assumed to cover all of in the absence of information on coverage. Post-PCV, there was a reduction in the maximum IPD incidence caused by a single serotype, as shown by datasets from Atlanta [192], Bogota [376], the Netherlands [384], and France [381]. As expected, VT, including VT10, had higher IPD incidences than NVT across all datasets and age groups pre-vaccination (Figure 4.5). These results highlight that variation in IPD burden may reflect genuine differences in disease, or artefacts of varying surveillance or reporting practices. However, given the heterogeneity of serotypes being carried between locations, it is also possible the higher disease burdens could result from the circulation of more invasive serotypes in carriage.
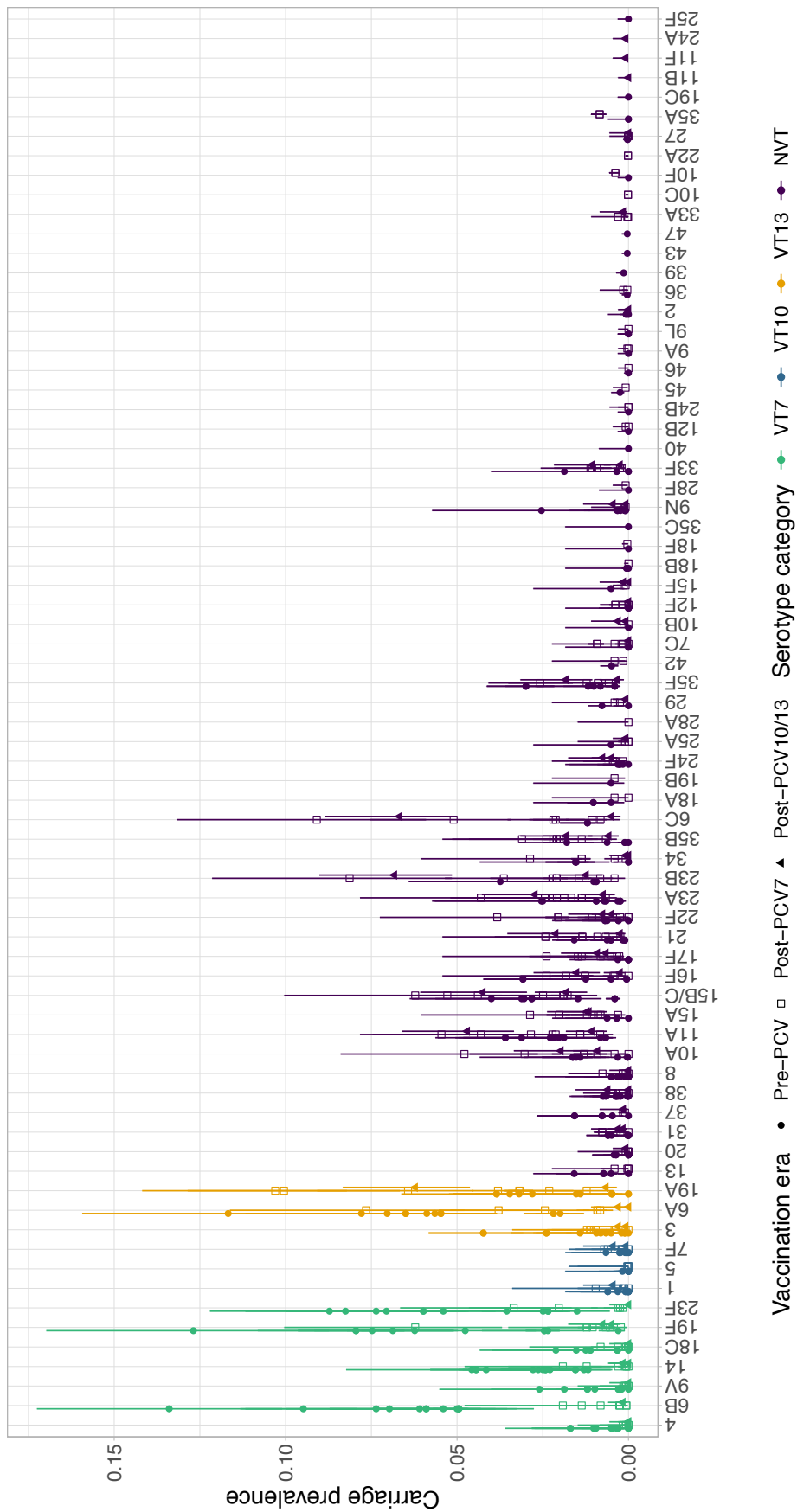
Figure 4.2: Carriage prevalence in children of each dataset by serotype. Each point represents the carriage prevalence of the respective serotype in a dataset. VT7: serotypes included in PCV7. VT10: additional serotypes included in PCV10. VT13: additional serotypes included in PCV13. NVT: serotypes not included in any PCV.
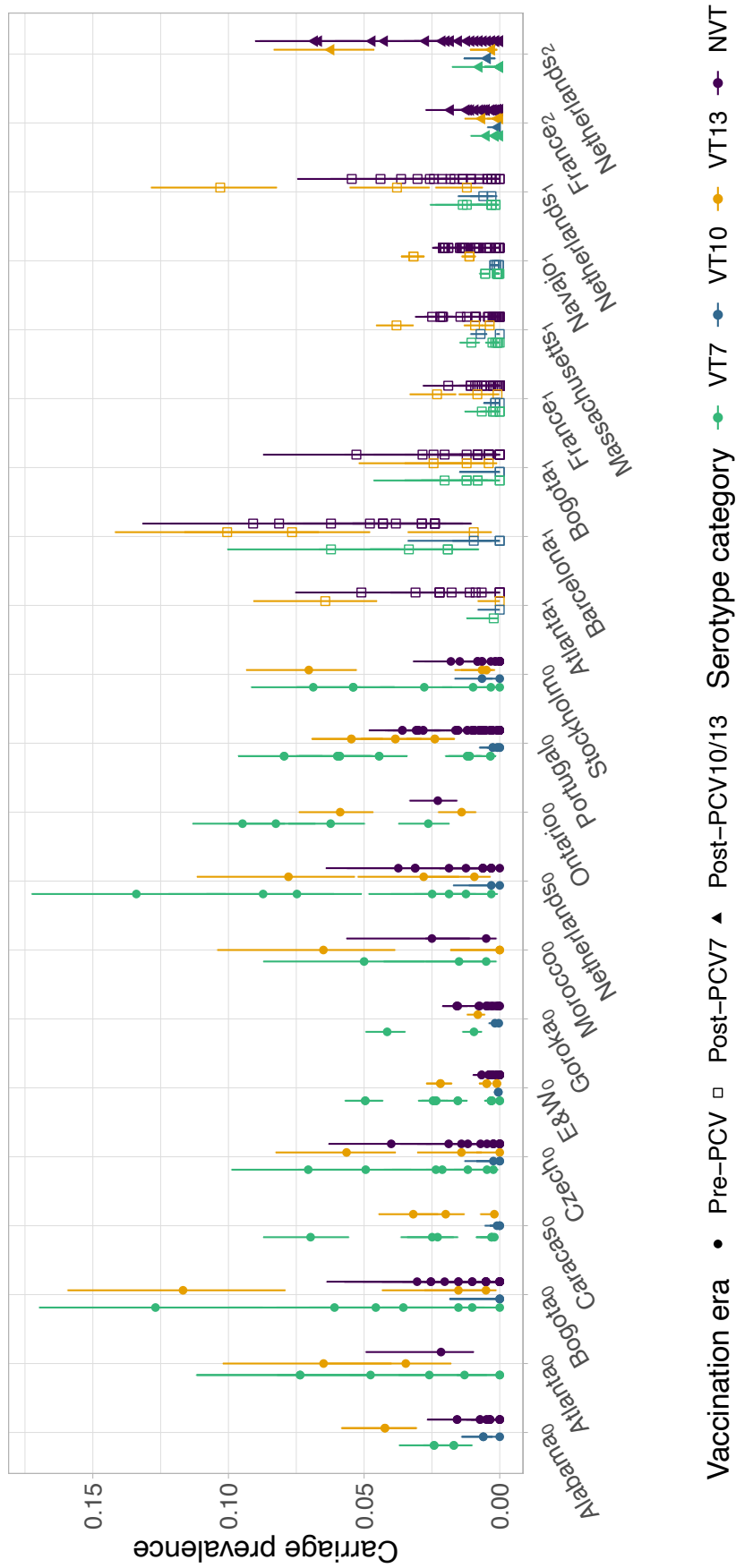
Figure 4.3: Carriage prevalence in children of each serotype by dataset. Each point represents the carriage prevalence of a serotype in the respective dataset. Subscript 0: pre-PCV; Subscript 1: post-PCV7; Subscript 2: post-PCV10/13.
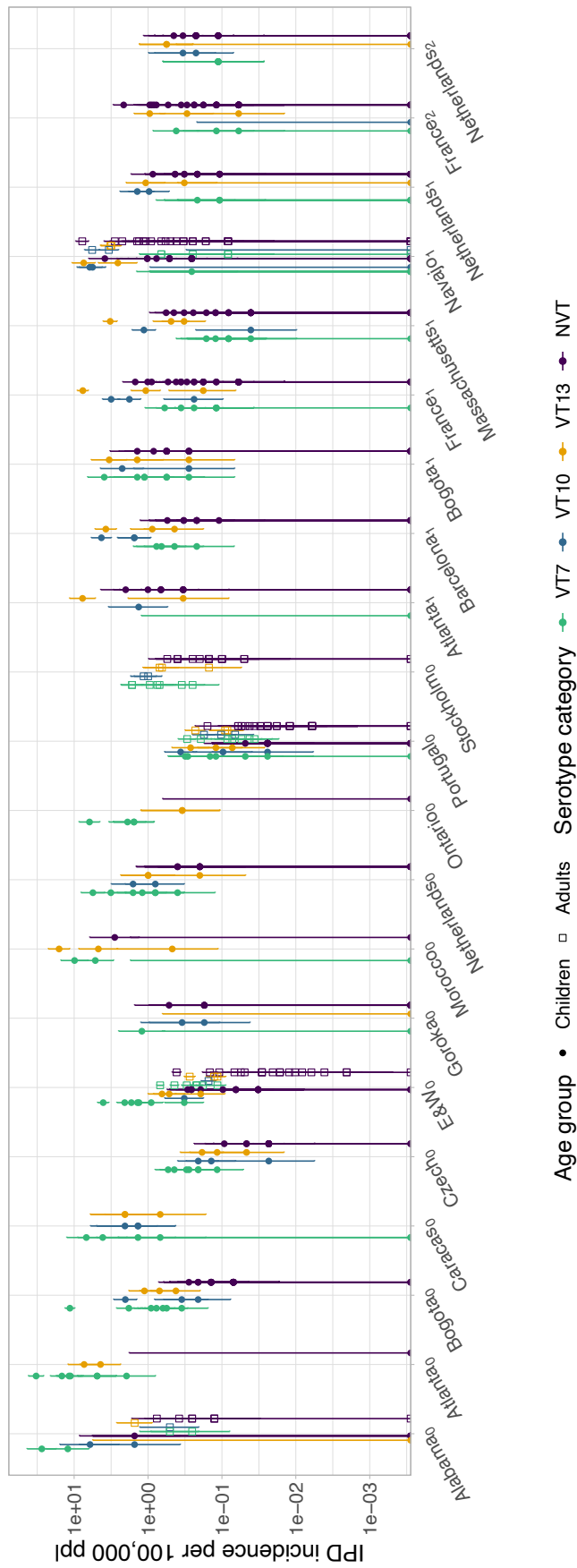
Figure 4.4: Invasive pneumococcal disease incidence of children ($<$ 18 years) and adults ($>$ 18 years) in each serotype by dataset. Subscript 0: pre-PCV; Subscript 1: post-PCV7; Subscript 2: post-PCV10/13.
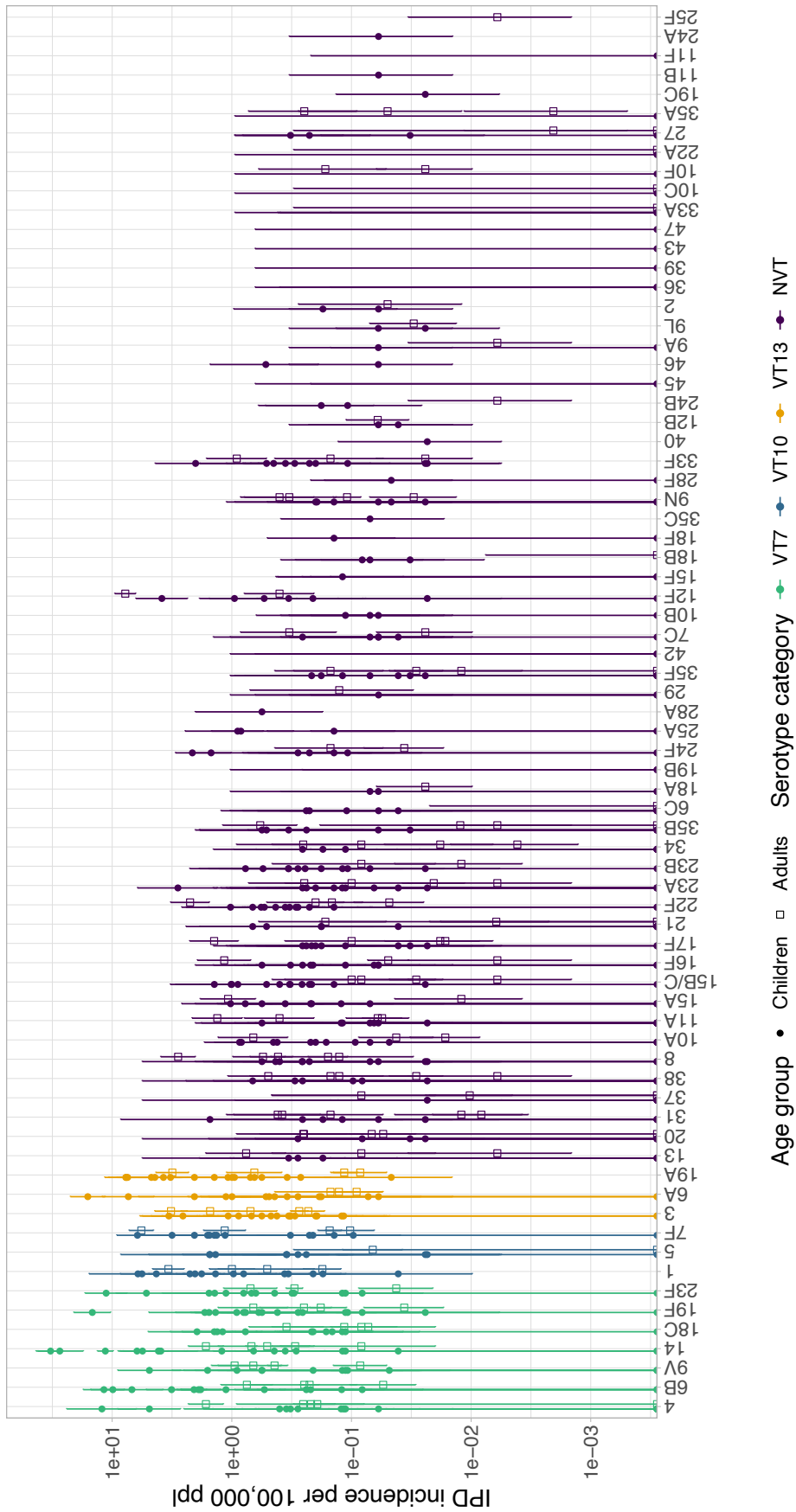
Figure 4.5: Invasive pneumococcal disease incidence of children ($<$ 18 years) and adults ($>$ 18 years) in each dataset by serotype.

4.3.2   Global invasiveness

Trends across serotypes

Each global invasiveness estimate is inferred from subsets of the overall dataset, correspond-
ing to those locations in which that serotype was observed (Table 4.4, Table 4.5). Hence greater
uncertainty (i.e. wider CrIs) was associated with those serotypes found in the lowest number
of datasets. Pre-PCV, global invasiveness values in children ranged from 10^{-5} (serotype 34)
to 3.6×10^{-3} (serotype 1) cases per carrier-year, with the majority of serotypes associated
with estimates in the range between 10^{-5} and 3×10^{-5} cases per carrier-year. Serotypes with
the highest global invasiveness in children were primarily VT, particularly VT10 serotypes
(Figure 4.6). NVT serotypes with high global invasiveness were serotypes 8, 9N, 12F, 17F,
18B, 20, and 22F. The VT7 serotypes 6B, 19F, and 23F and VT13 serotypes, had similar or
lower invasiveness relative to many of these NVT. Post-PCV, invasiveness values in children
ranged from 10^{-5} (serotype 33A) to 1.7×10^{-2} (serotype 1) cases per carrier-year, with the
majority of serotypes associated with estimates in the range between 10^{-4} and 2×10^{-4} cases
per carrier-year. VT10 serotypes were also high in invasiveness post-PCV, whereas NVT
serotype 12F's invasiveness was significantly higher than most other serotypes (Figure 4.7),
although this analysis represents a smaller and different combination of datasets than the
pre-PCV analysis.

The model was also fitted to data from adult IPD, relative to carriage in infants. Pre-PCV,
global invasiveness values in adults ranged from 1.8×10^{-5} (serotype 21) to 2.4×10^{-3}
(serotype 8) cases per carrier-year, with the majority of serotypes associated with estimates
in the range between 2.5×10^{-5} and 5×10^{-5} cases per carrier-year. NVT serotypes tended
to have higher invasiveness compared to VT serotypes (Figure 4.6). Serotypes 8, 9N, 13,
17F, 20, 22F and 38 were some of the NVTs with higher invasiveness, many of which have
subsequently been associated with serotype replacement in adult IPD. Serotype 3 was the
only VT13 serotype among the top 10 most invasive serotypes, reflecting its prevalence in
disease post-vaccination with PCV7. Regardless, serotypes 8, 9N, 22F and 38 may be key
serotypes for both children and adults, particularly as they have not yet been included in any

of the licensed PCV formulations.

Many of the serotypes do not exhibit a significant difference between the overall invasiveness estimates in children and adults (Figure 4.6), although these estimates are derived from a different combination of datasets. The exception in NVT is serotype 8, which has a significantly higher invasiveness in adults. All of the VT serotypes with significant difference between the two age groups were unsurprisingly significantly higher in children than adults, as these (19F, 23F, 6B) are known as the paediatric serotypes because they were common in infant disease. The exception was VT13 serotype 3 which was significantly higher in adults. The VT serotypes 1, 4 and 9V did not exhibit significantly different invasiveness between children and adults. Therefore, global invasiveness is not always significantly different in serotypes across age groups.
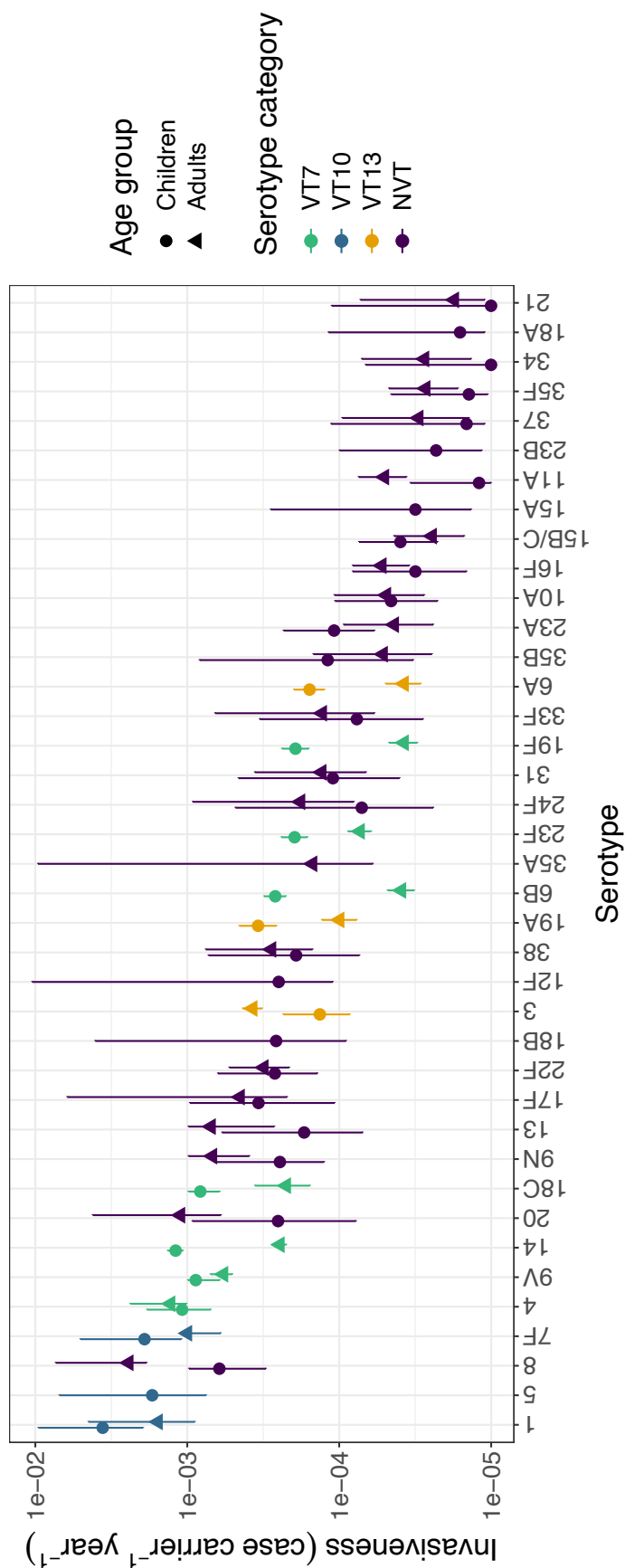
Figure 4.6: Global invasiveness in children and adults pre-PCV of serotypes included in more than two datasets. Error bars represent 95% credible intervals.
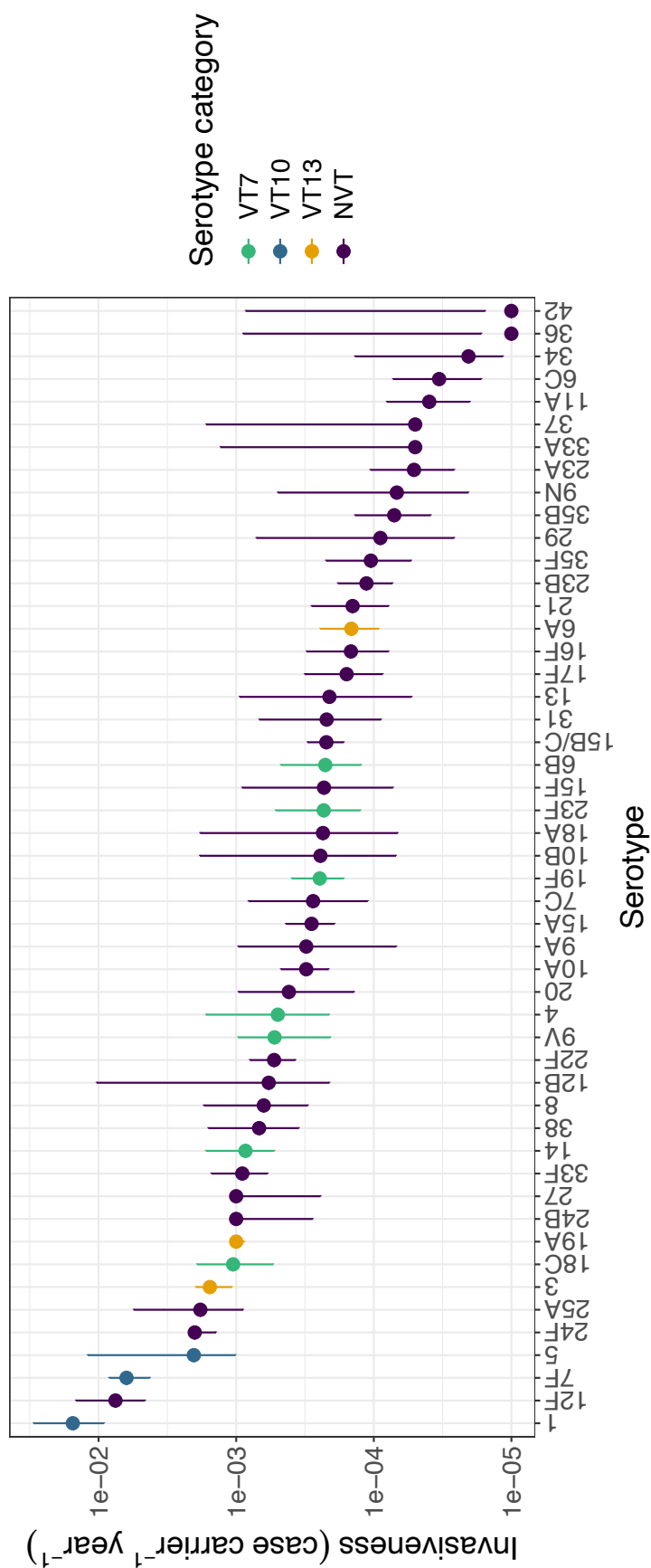
Figure 4.7: Global invasiveness in children post-PCV7 of serotypes included in more than two datasets. Error bars represent 95% credible intervals.

Genetic background and serotype correlation in children and adults

To test whether these correlations exist across a diversity of genetic backgrounds, serotype invasiveness pre-PCV was plotted against the number of strains (Global Pneumococcal Sequence Clusters, GPSCs) in which the serotype was expressed in the GPS collection [351]. The number of strains was not correlated with global invasiveness in children (Figure 4.8), but had a statistically significant negative correlation in adults (Figure 4.9, $p < 0.05$). If serotype does not cause differences in invasiveness, then serotypes found in more strains should have non-extreme invasiveness values. The serotypes with the highest overall invasiveness in children (1, 5 and 7F) were all found in fewer than five GPSCs, despite each being well-represented in the isolates from IPD. This means it is hard to ascertain whether invasiveness is attributable to serotype. Serotype 8, which had the highest overall invasiveness estimate in adults pre-PCV was found in fewer than 10 strains globally. VT serotypes with slightly lower invasiveness (23F, 19F, 19A, 6B and 6A) in children and adults were reported to have more than 30 strains. These are the paediatric serotypes that are common in infant carriage and disease. In infants, they have middling invasiveness, as might be expected if invasiveness were driven by other loci, and the estimated value is assumed to be an average across strains with different levels of invasiveness. Hence results are consistent with non-serotype loci contributing to invasiveness, as many other variable loci will likely be identified to contribute to differences in invasiveness.
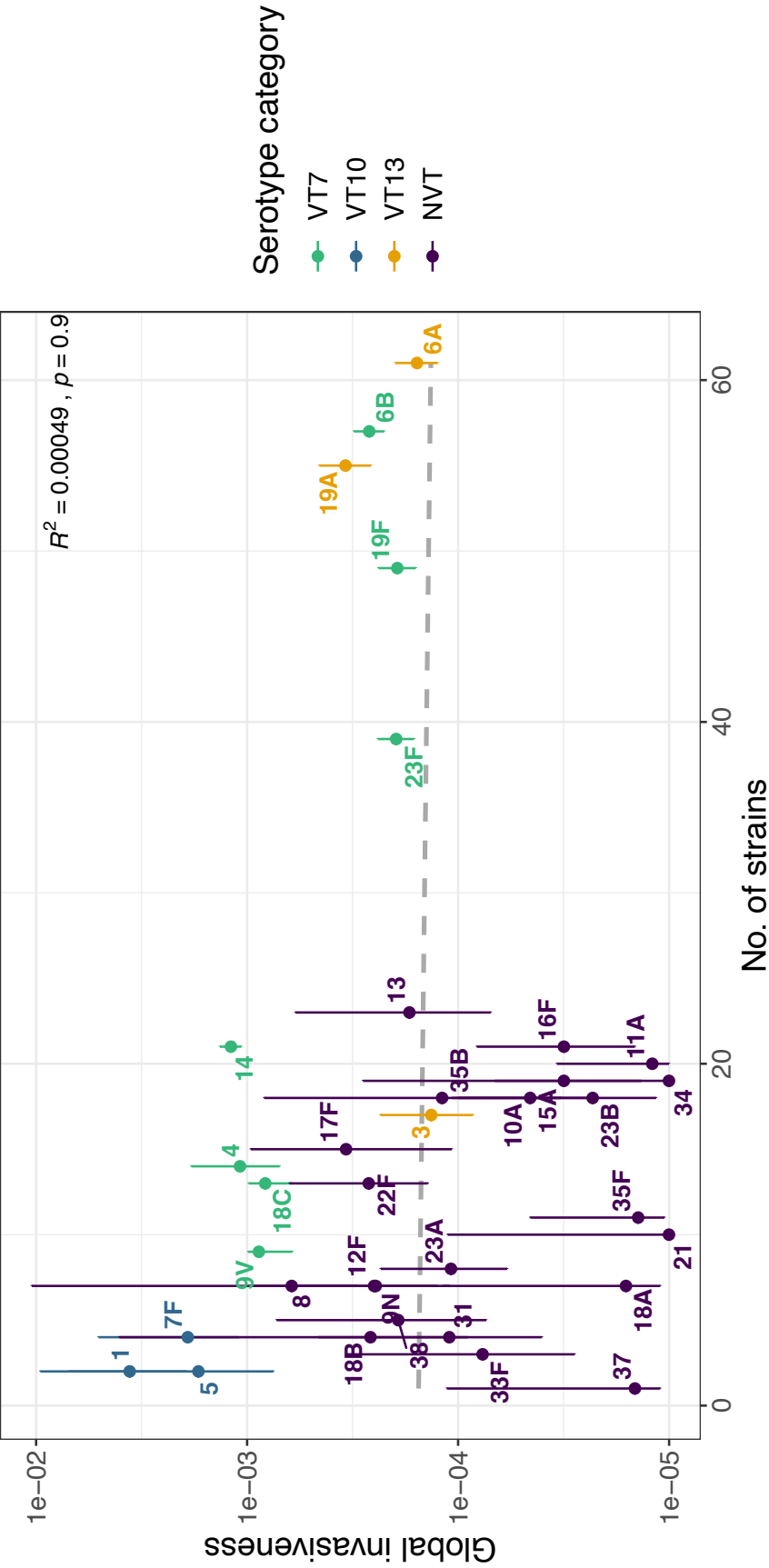
Figure 4.8: Comparison of global invasiveness in children and number of strains from global pneumococcal sequence cluster (GPSC).
Error bars represent 95% credible intervals.
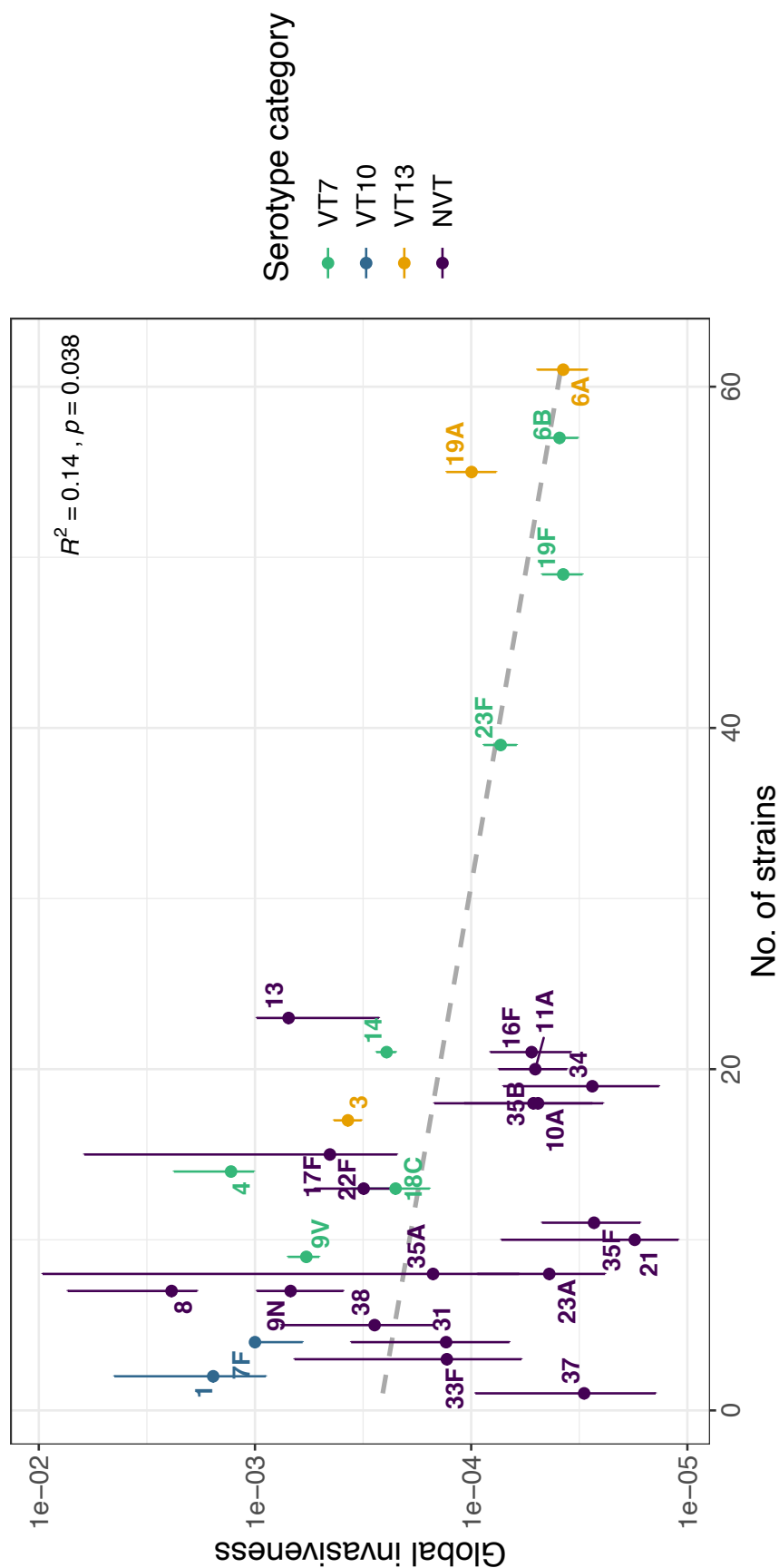
Figure 4.9: Comparison of global invasiveness in adults and number of strains from global pneumococcal sequence cluster (GPSC). Error bars represent 95% credible intervals.

4.3.3 Local invasiveness

As invasiveness may vary between strains of the same serotype, the local invasiveness was also estimated as the serotype-specific invasiveness within each dataset to investigate this heterogeneity. These rates showed a large range of estimates for each serotype for both children and adults (Figure 4.10). Maximum a posteriori invasiveness estimates at the bottom of the 95% CrI represent the serotypes that had zero cases of carriage and a non-zero but low number of disease cases. The maximum a posteriori estimates of serotype local invasiveness within datasets (Figure 4.11) ranged from between 7.0×10^{-5} and 2.3×10^{-3} (Ontario pre-PCV [375]) to between 7.0×10^{-5} and 3.7×10^{-2} (Goroka pre-PCV [342]). Ultimately there was no serotype-specific trend in invasiveness estimates across all datasets, as serotypes that were highly invasive in some datasets were low in others (Table 4.6, Table 4.7). For example, in children pre-PCV, serotype 8 was highly invasive in England and Wales but less invasive in Morocco; 9V was highly invasive in the Czech Republic but less invasive in Caracas; 19A was highly invasive in Morocco but less invasive in Atlanta; and 31 was highly invasive in Alabama but less invasive in Portugal. Post-PCV7, serotype 4 was highly invasive in Massachusetts but less invasive in France. Post-PCV10/13, serotype 38 was highly invasive in France but less invasive in the Netherlands. In adults pre-PCV, serotype 8 was highly invasive in England and Wales but less invasive in Alabama; 4 was highly invasive in Stockholm but less invasive in Alabama and Navajo post-PCV7; and 31 was highly invasive in Alabama but less invasive in Portugal. Around two thirds of the serotypes in both age groups had cases in which they were not detected in carriage but caused disease, suggesting they were rarely present in the nasopharynx but sufficiently invasive to be found in IPD [65, 66, 401], although this may reflect different carriage and IPD sample sizes.

Despite the small sample size of datasets including both children and adults (Figure 4.12), 19A appears to be significantly more invasive in children than adults (England and Wales (E&W) [380] and Navajo [223, 379]). In the E&W study, VT7 serotypes 6B, 18C, 19F and 23F as well as 15B/C, were also significantly more invasive in children than adults. In the Alabama study [373], only serotypes 4 and 14 were significantly more invasive in children than adults.

In the Portugal study [352], there was no significant difference between invasiveness of serotypes carried by both children and adults. All of the studies that included both children and adults had respectively similar invasiveness ranges for both age groups. These are directly correlated to IPD incidence, since carriage proportions were considered the same for children and adults. IPD incidence in children is higher than in adults, however in the Navajo population they may be similar because they are known to be a population with high IPD incidence due to comorbidities [402]. The E&W dataset stratified age by less than and greater than 5 years old, and the addition of children between the ages of 5 and 18 may have raised the range of the disease incidence and thus invasiveness estimates in adults.

Only four studies had both pre- and post-vaccination data (Figure 4.13). Studies conducted in Atlanta [192] and Bogota [376] were pre- and post-PCV7, with the one in Atlanta reporting only three serotypes in common pre- and post-vaccination and none significantly different in invasiveness. The Bogota study had many of the same serotypes present post-vaccination but these did not show a clear trend in change of invasiveness in the serotypes detected in both vaccination periods. The third study, conducted in France pre- and post-PCV13 did not show any clear trend in invasiveness either [381]. The last study, conducted in the Netherlands with pre-PCV, post-PCV7 and post-PCV10 data, showed an increase in invasiveness in serotype 3 and a decrease in serotype 7F, though non-significant. Many of the common serotypes in this study had non-significant differences across vaccination periods. This suggests that invasiveness does not change with vaccination as no serotypes (both VT and NVT) had a significant difference in invasiveness post-vaccination regardless of vaccine, but it merits further exploration when additional pre- and post-vaccination datasets from the same setting become available, since there were only six datasets post-PCV7 and one post-PCV13 and only four settings with both pre- and post-vaccination data. Therefore, vaccination did not appear to substantially change the invasiveness associated with serotypes, relative to either the variation observed between serotypes, or the same serotype in different locations.
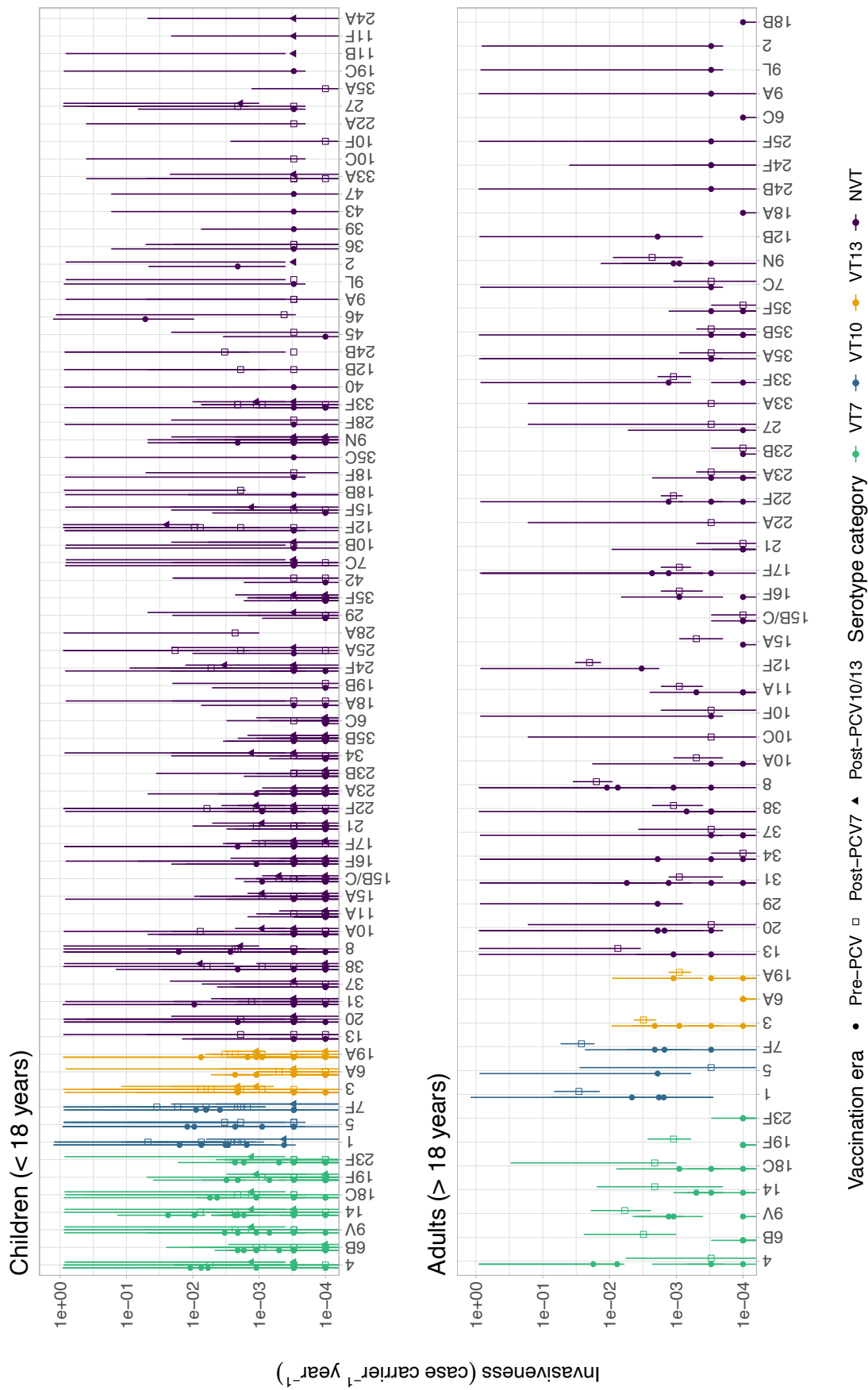
Figure 4.10: Local invasiveness of each serotype in children and adults by serotype. Each point represents a dataset which includes the respective serotype. Error bars represent the 95% credible interval.

Table 4.6: Serotypes ranked by local invasiveness (most to least invasive) for each dataset reporting children. VT7: serotypes included in PCV7. VT10: serotypes included in PCV10. VT13: serotypes included in PCV13. NVT: serotypes not included in currently licensed vaccines.

Pre-PCV											Post-PCV7							Post-PCV10/13	
Alabama	Atlanta	Bogota	Caracas	Czech	Eng&Wales	Goroka	Morocco	Netherlands	Ontario	Portugal	Atlanta	Barcelona	Bogota	France	Massachusetts	Navajo	Netherlands	France	Netherlands
31	14	1	1	9V	8	46	19A	7F	14	1	7F	1	3	25A	4	1	27	12F	27
5	4	14	5	1	4	1	23F	8	19A	20	22F	5	1	7F	18C	7F	7F	3	8
14	18C	9V	7F	12F	18C	5	6A	14	19F	5	38	7F	28A	12F	33F	12F	1	38	14
1	19F	5	14	17F	7F	2	6B	1	23F	9L	33F	3	8	1	12B	19A	24B	24F	18C
4	9V	10B	19F	19A	9V	14	23A	18C	6B	19C	12F	14	9V	24F	18B	3	10A	11B	23F
13	23F	12F	3	28F	14	31	3	4	11A	7F	13	19A	14	10A	20	8	12F	15F	34
20	6B	15A	6A	33F	17F	36	4	9V	6A	8	3	10A	25A	14	1	15A	15A	2	4
38	19A	18B	6B	40	20	39	8	10A		9N	19A	11A	18C	5	7F	19F	16F	4	9V
8	6A	18F	19A	5	3	43		17F		10A	21	15A	19A	19A	19A	33F	17F	6A	3
3	11A	22F	23F	14	38	47		19A		11A	15B/C	15B/C	6B	46	22F	38	18C	7C	1
37		24F	4	4	9N	13		23A		14	19F	16F	20	10B	3	10C	22F	10A	10A
		35C	9V	7F	19A	20		3		15A	23B	17F	22F	16F	6A	14	24F	14	10B
		7C		9N	16F	21		6B		15B/C	11A	19F	4	18A	10A	15B/C	3	15A	15F
		7F		10A	15B/C	29		9N		16F	15A	21	5	18C	13	16F	31	19A	22F
		8		11A	22F	3		11A		18A	23A	22F	11A	24B	14	17F	36	19F	24F
		13		15B/C	19F	34		15A		18C	35B	23A	15A	31	17F	18C	4	21	31
		16F		18C	23F	38		15B/C		19A	6C	23B	15B/C	7C	18F	20	42	22F	33A
		18A		19F	6B	4		16F		19F		23F	16F	9A	21	21	8	24A	33F
		18C		23A	10A	42		19F		21		34	19F	9L	23F	22A	9V	29	37
		19A		23F	18B	45		21		22F		35B	23B	9V	33A	22F	11A	33F	6A
		25A		3	23A	8		22F		23A		6A	23F	12B	36	23A	14	10B	7F
		3		35F	27			23B		23B		6B	24F	17F	38	23B	15B/C	11A	11A
		4		37	31			23F		23F		6C	35B	19F	6B	27	19F	11F	15A
		6B		38	35B			24F		24F			10A	22F	7C	31	21	15B/C	15B/C
		10A		6A	6A			33F		3			13	23F	9A	33A	23A	16F	16F
		11A		6B	11A			35B		31			17F	33F	9V	34	23B	17F	17F
		15B/C		8	21			38		33F			18A	6B	11A	35B	23F	20	19A
		15F			35F			6A		34			19B		15A	37		23A	19F

	Pre-PCV											Post-PCV7							Post-PCV10/13	
	Alabama	Atlanta	Bogota	Caracas	Czech	Eng&Wales	Goroka	Morocco	Netherlands	Ontario	Portugal	Atlanta	Barcelona	Bogota	France	Massachusetts	Navajo	Netherlands	France	Netherlands
			19B								35F			23A	8	15B/C	4	33A	23B	21
			19F								37			29	9N	15F	5	33F	25A	23A
			23A								4			35F	11A	16F	6B	35B	31	23B
			23B								6A			42	15A	19F	7C	35F	34	35B
			23F								6B			6A	15B/C	23A	9N	38	35B	35F
			34								6C			6C	15F	23B	9V	6A	35F	38
			35F								9V			7C	20	25A	10A	6B	6B	6C
			6A												21	31	10F	6C	6C	9N
			9N												23A	34	11A	9N	7F	
														23B	35B	35A		9N	
														28F	35F	35F			
														29	37				
														34	6C				
														35B	9N				
														35F					
														4					
														45					
														6A					
														6C					

Table 4.7: Serotypes ranked by local invasiveness (most to least) for each dataset reporting adults. VT7: serotypes included in PCV7. VT10: serotypes included in PCV10. VT13: serotypes included in PCV13. NVT: serotypes not included in currently licensed vaccines.

	Pre-PCV			Post-PCV7
Alabama	**England&Wales**	**Portugal**	**Stockholm**	**Navajo**
31	8	12B	4	1
29	4	5	8	7F
34	17F	1	1	12F
1	3	20	12F	8
13	7F	8	20	13
14	20	10F	7F	9V
20	9V	13	17F	3
3	16F	17F	22F	6B
38	9N	24B	31	9N
8	38	25F	33F	14
37	14	35B	19A	18C
4	34	38	9N	19F
	35A	7C	9V	22F
	37	9A	18C	33F
	10A	9L	3	38
	18C	24F	11A	11A
	19A	4	14	16F
	22F	7F	2	17F
	31	9N	35A	19A
	11A	10A	23A	31
	15B/C	11A	24F	10A
	18B	14	35F	15A
	19F	15A	38	10C
	21	15B/C	15B/C	10F
	23A	16F	19F	20
	23F	18A	21	22A
	27	18C	23F	23A
	35B	19A	35B	27
	35F	19F	6A	33A
	6A	21	6B	35A
	6B	22F		35B
		23A		37
		23B		4
		23F		5
		3		7C
		31		15B/C
		33F		21
		34		23B
		35F		34
		37		35F
		6A		
		6B		
		6C		
		9V		
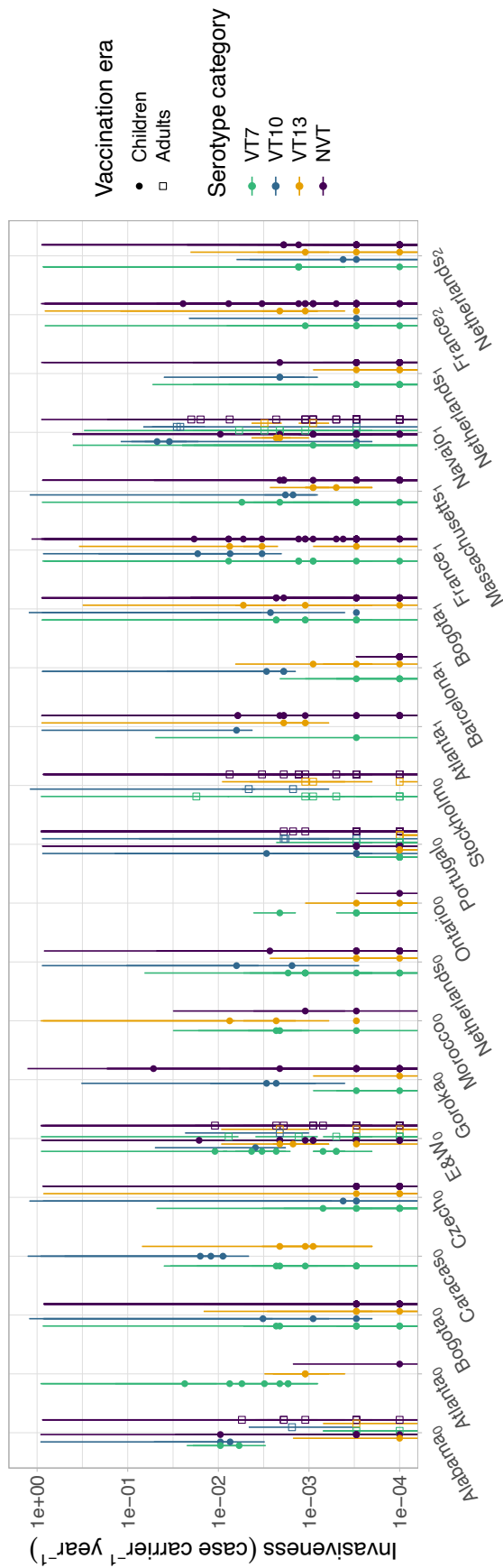
Figure 4.11: Local invasiveness of each serotype in children and adults by dataset. Each point represents a serotype in the respective dataset. Error bars represent the 95% credible interval. Subscript 0: pre-PCV; Subscript 1: post-PCV7; Subscript 2: post-PCV10/13.
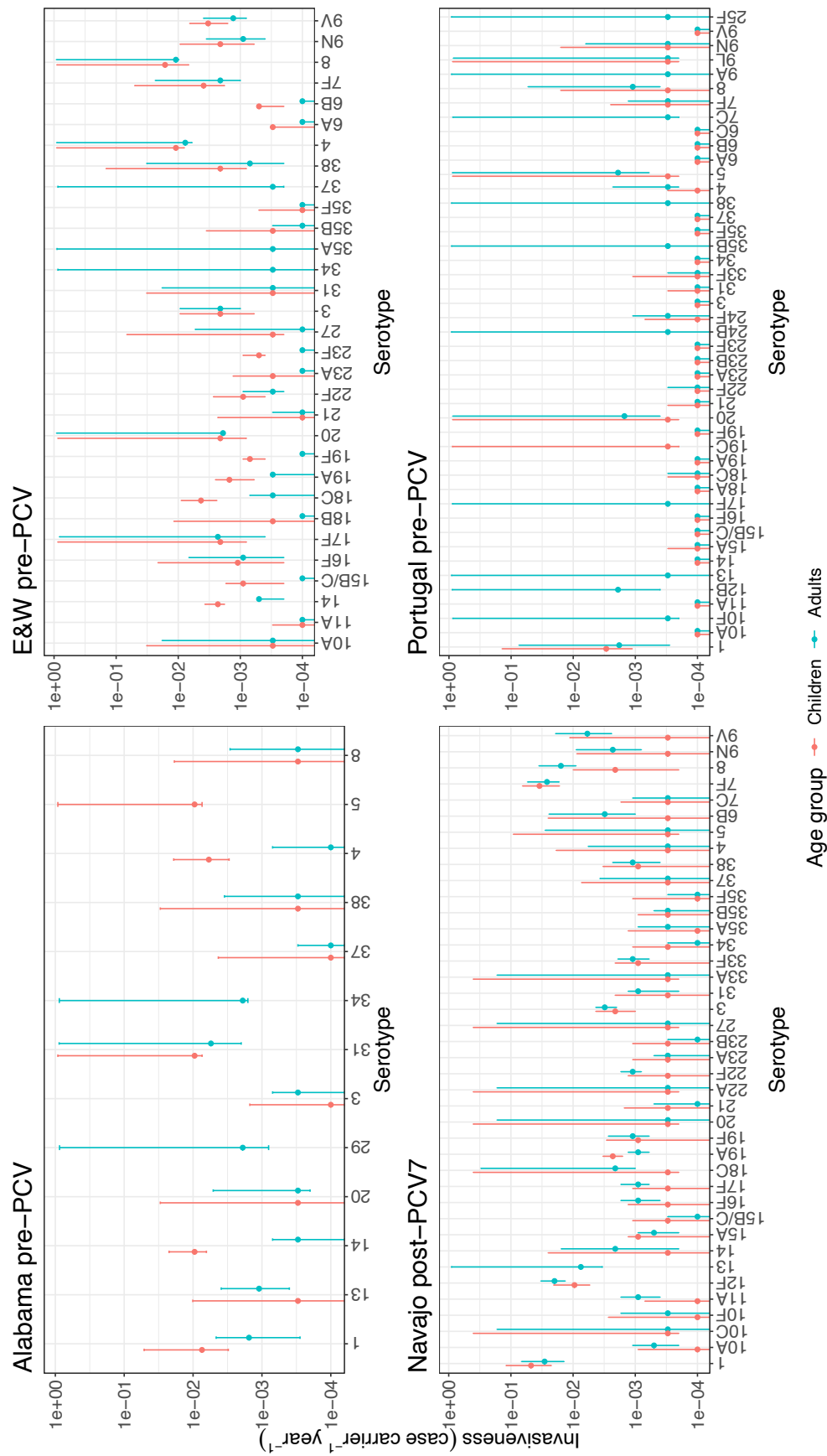
Figure 4.12: Local invasiveness of serotypes in datasets that included data on both children and adults. Error bars represent the 95% credible interval.
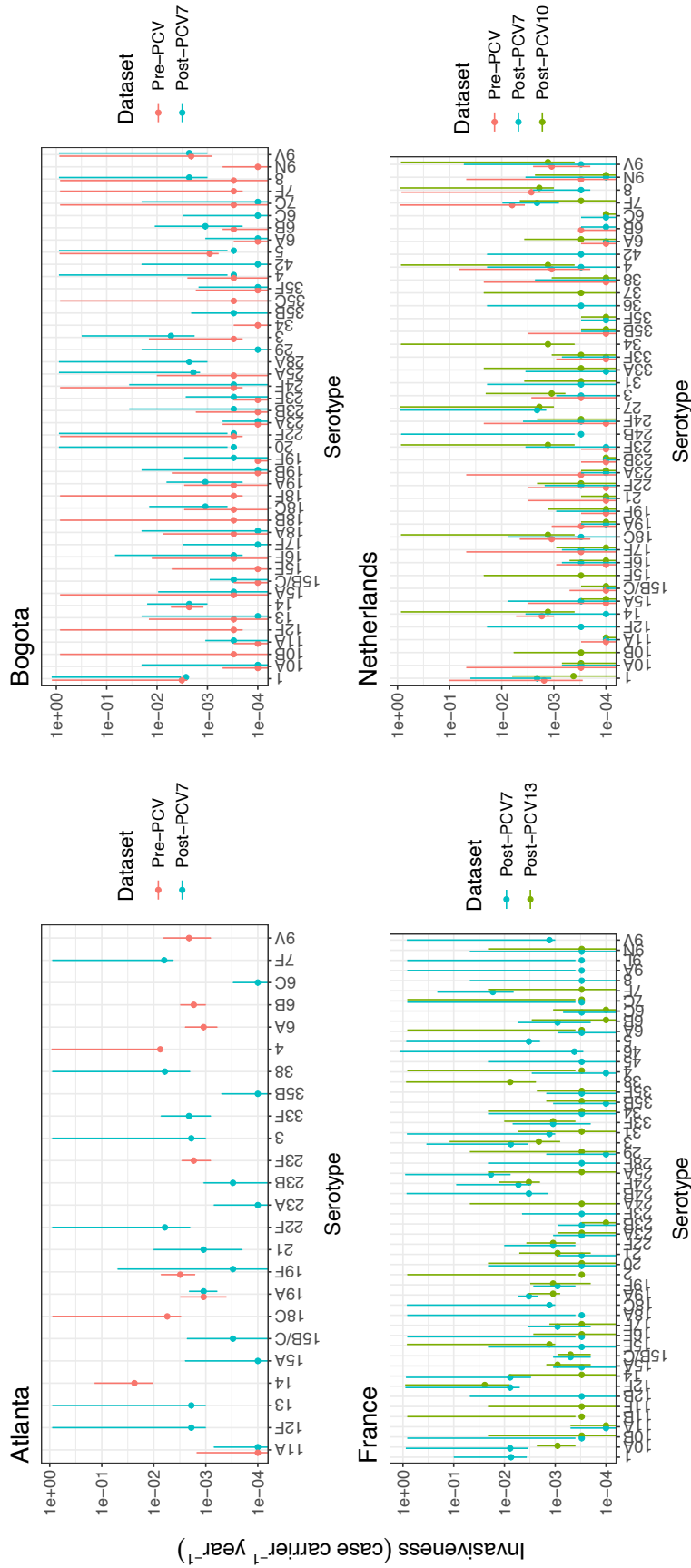
Figure 4.13: Local invasiveness of serotypes in datasets that included pre- and post-vaccination periods. Error bars represent the 95% credible interval.

4.3.4 Model comparison

Datasets did not have consistently higher or lower local invasiveness measures for serotypes compared to the global invasiveness, indicating that location may not be the only factor impacting a serotype's "true" invasiveness (Figure 4.14).

To formally test the global invasiveness and local invasiveness models for all serotypes, the relative evidence of the two was estimated using BF. BF showed evidence for the global invasiveness model over the local invasiveness model in most serotypes in children, showing stronger evidence for the global model as the number of datasets increased (Figure 4.15), with the exception of serotype 14 which showed extreme evidence for the local invasiveness model. In adults, the BF also favoured the global model over the local in most serotypes, though more weakly than in the case of children. The exception was serotype 16F, for which there were only two datasets and which showed extreme evidence for the local model. Taken together, these results are not surprising since the global invasiveness model was constrained to one invasiveness estimate across all datasets, making the number of parameters used $number$ $of\ datasets - 1$ fewer than its local model counterpart. The maximum number of datasets for a serotype was 11, and whether the preference for the global model will hold for a larger number of datasets is unknown as of yet. The effect may also work in counterbalance when the variance between few datasets for a serotypes is large, such as for serotype 14 in children and serotype 16F in adults. The evidence ratio suggests that the global model is preferable to estimating invasiveness compared to the local model as the number of datasets included for a serotype increases. Hence only those serotypes for which most data are available have the power to reject the local invasiveness model.
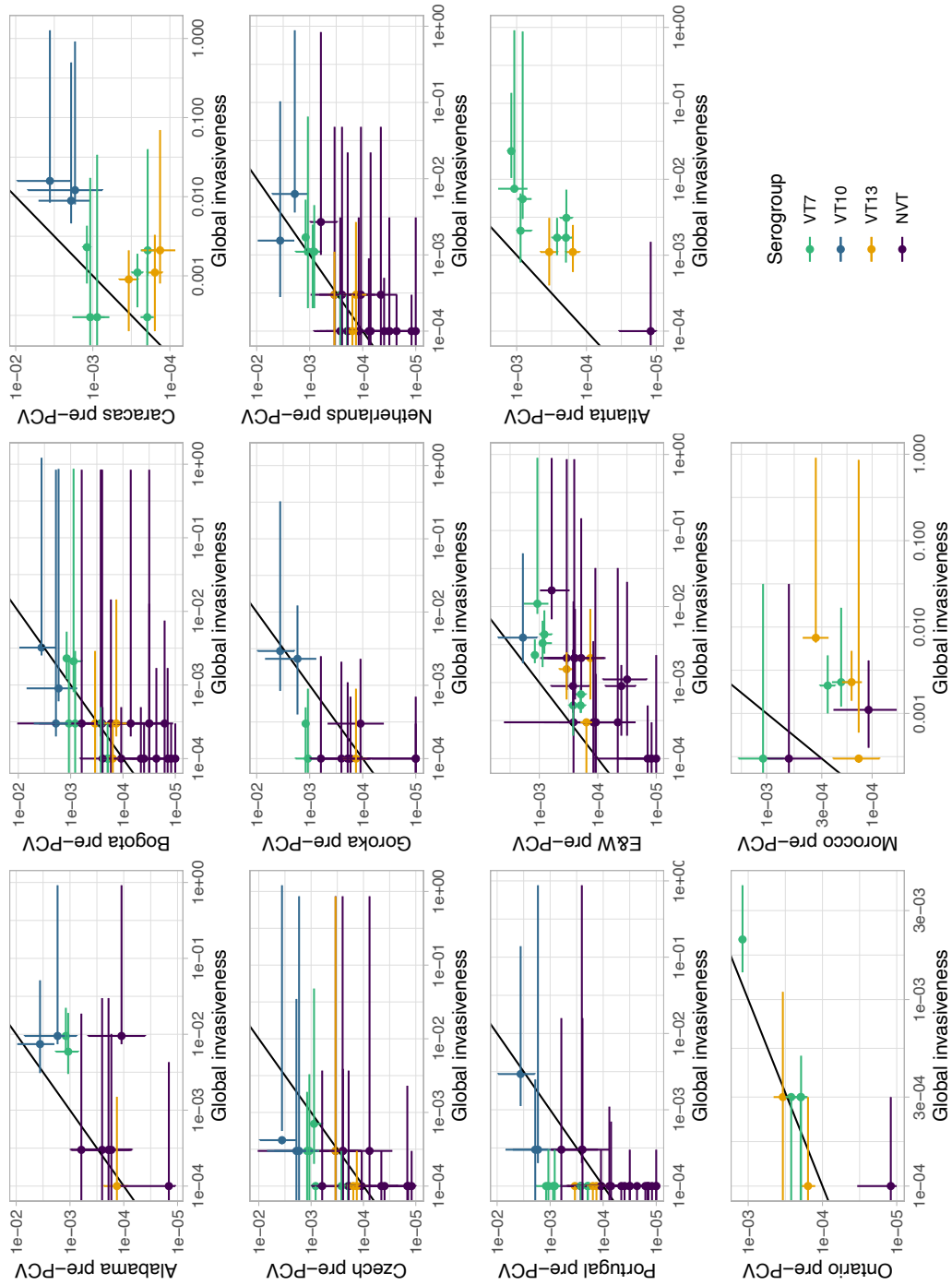
Figure 4.14: Comparison of pre-PCV study-specific local invasiveness and global invasiveness. Error bars represent the 95% credible interval. E&W: England and Wales.
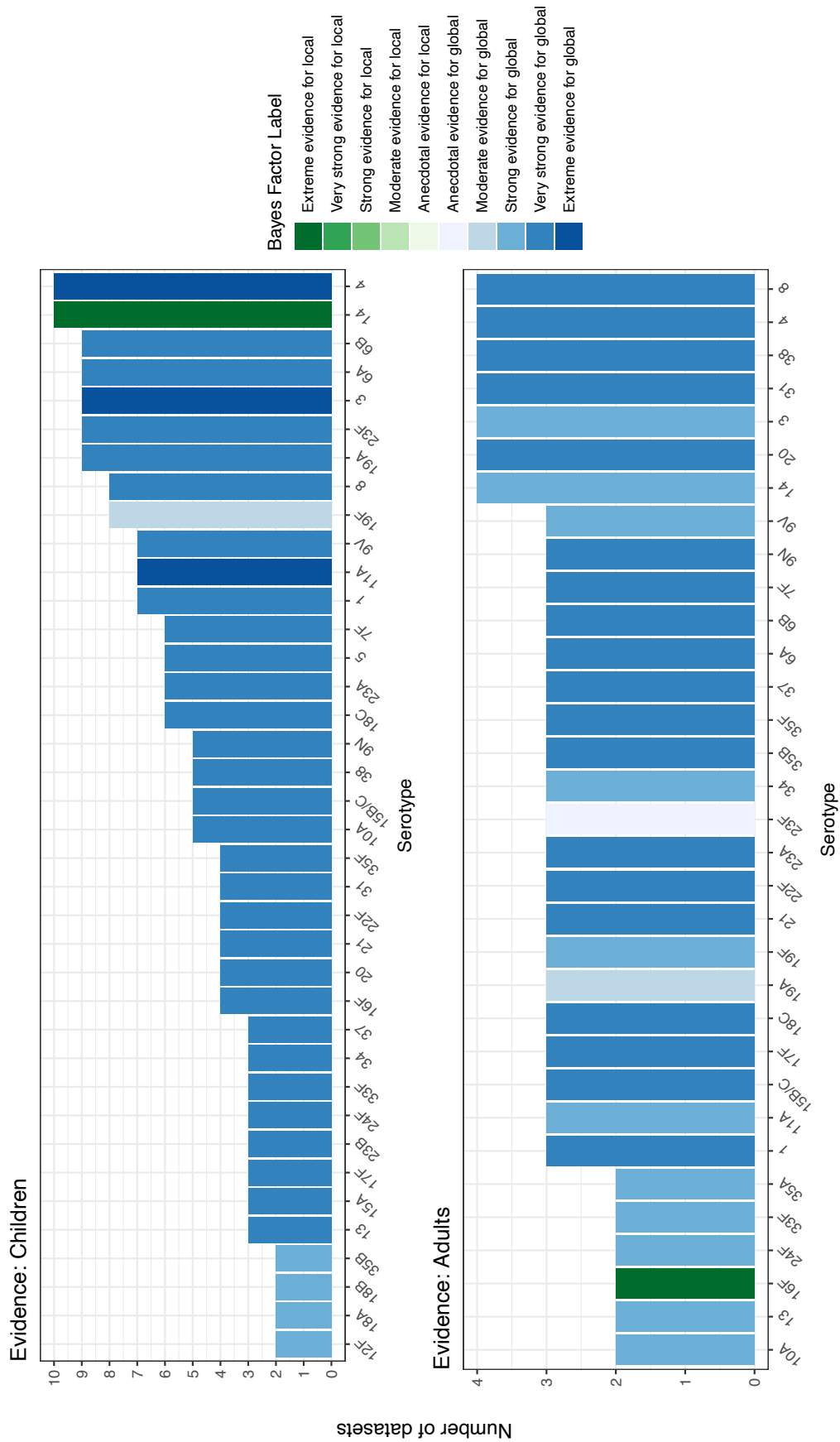
Figure 4.15: Pre-PCV model comparison of evidence in children and adults vs the number of datasets including the respective serotypes.

This becomes further evident when comparing the posterior distributions of global and local invasiveness of serotypes in children (Figure 4.16). As the number of datasets increases, the posterior distribution of the global invasiveness becomes narrower because the distribution becomes more 'informed'. This suggests that while a single global invasiveness is preferred to an amalgamation of local invasiveness estimates, a universal invasiveness will not always be appropriate for describing each location individually. In particular, the posterior distributions of local invasiveness of serotype 14, which dominated the serotype landscape pre-PCV, show variability between datasets. This is inconsistent with the posterior distributions of the local invasiveness of serotype 4, which are almost the same combination of datasets as those of serotype 14, but that have similar local invasiveness posterior distributions. In other words, global invasiveness does not capture all serotypes' invasiveness in datasets equally well, and this is a caveat of using a universal invasiveness measure for each serotype.

Using the local invasiveness method, we can directly compare the probability of serotypes causing disease given carriage between different serotypes and between the same serotype in different datasets. Taking for example VT7 serotypes 4 and 14 and epidemic serotypes 1 and 5 suggests that the variation in invasiveness is driven by location or location-specific loci and cannot be ascribed solely to serotypes (Figure 4.17). The VT7 serotypes which are more prevalent in carriage and disease have smaller CrIs of local invasiveness and are sometimes significantly different between datasets, as well as from the global invasiveness. On the other hand, the rarer epidemic serotypes 1 and 5 have much wider intervals and are only sometimes significantly different from the global invasiveness, as is the case for serotype 1 in Goroka. This indicates that the global invasiveness may provide a reasonable estimate in the absence of other data for the rarer serotypes.

4.3.5   Comparison with previous results

Comparing the posterior distributions of invasiveness in children with those of Weinberger et al. [223], these local invasiveness distributions differ sporadically from theirs depending on the dataset, whereas the global invasiveness distribution tends to overlap with their results

176

(Figure 4.18). These global invasiveness estimates included the dataset from the population
on which Weinberger et al.'s analysis was based (Navajo post-PCV7). Interestingly, the
local invasiveness estimates for the Navajo dataset were slightly higher than those found
in Weinberger et al.'s paper. The narrowness of the Weinberger et al. posterior distribution
makes it undesirable as a prior, and it likely reflects the plenitude of data used from multiple
time periods compared to the one used for the analysis in this chapter.
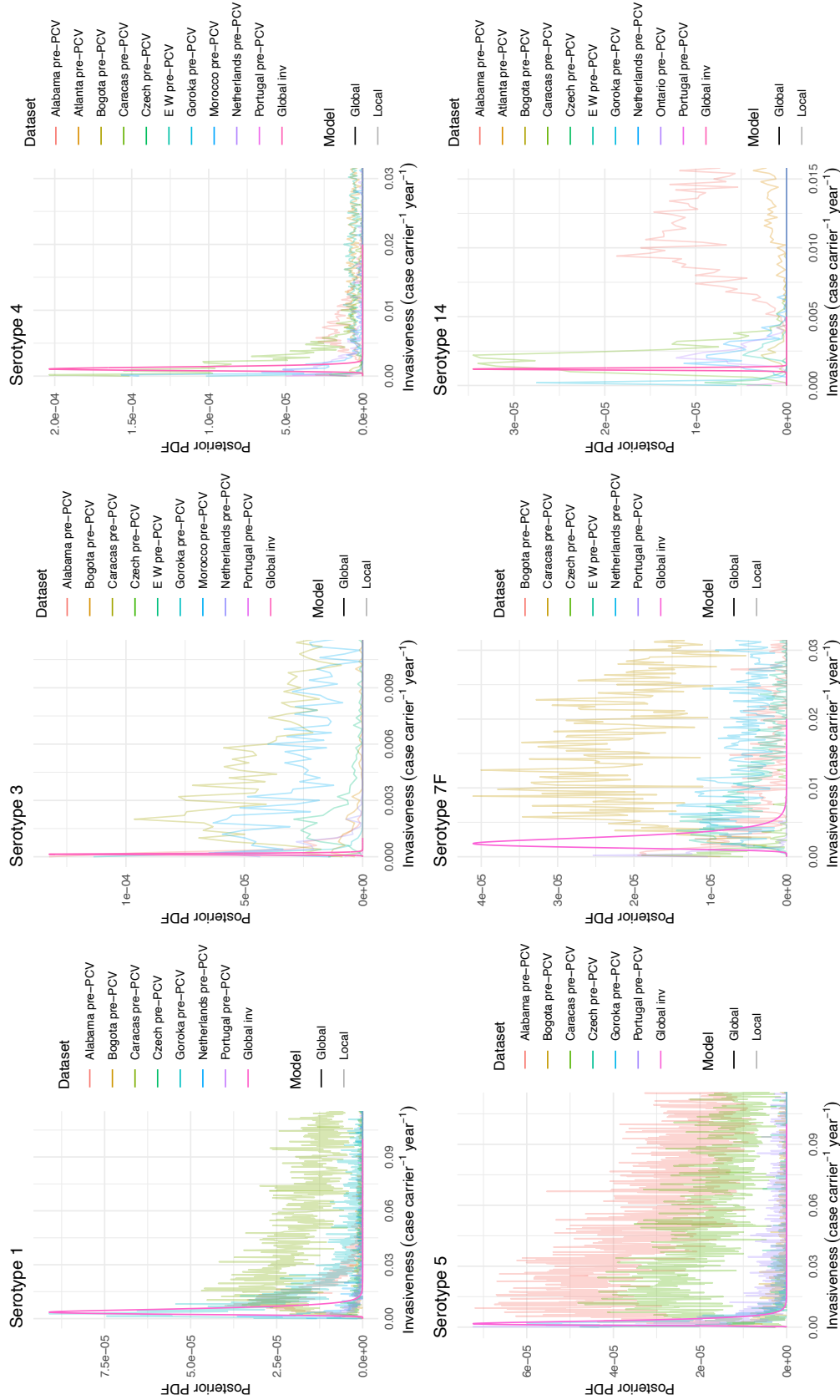
Figure 4.16: Comparison of posterior distributions of study-specific local invasiveness with global invasiveness. EW: England and Wales.
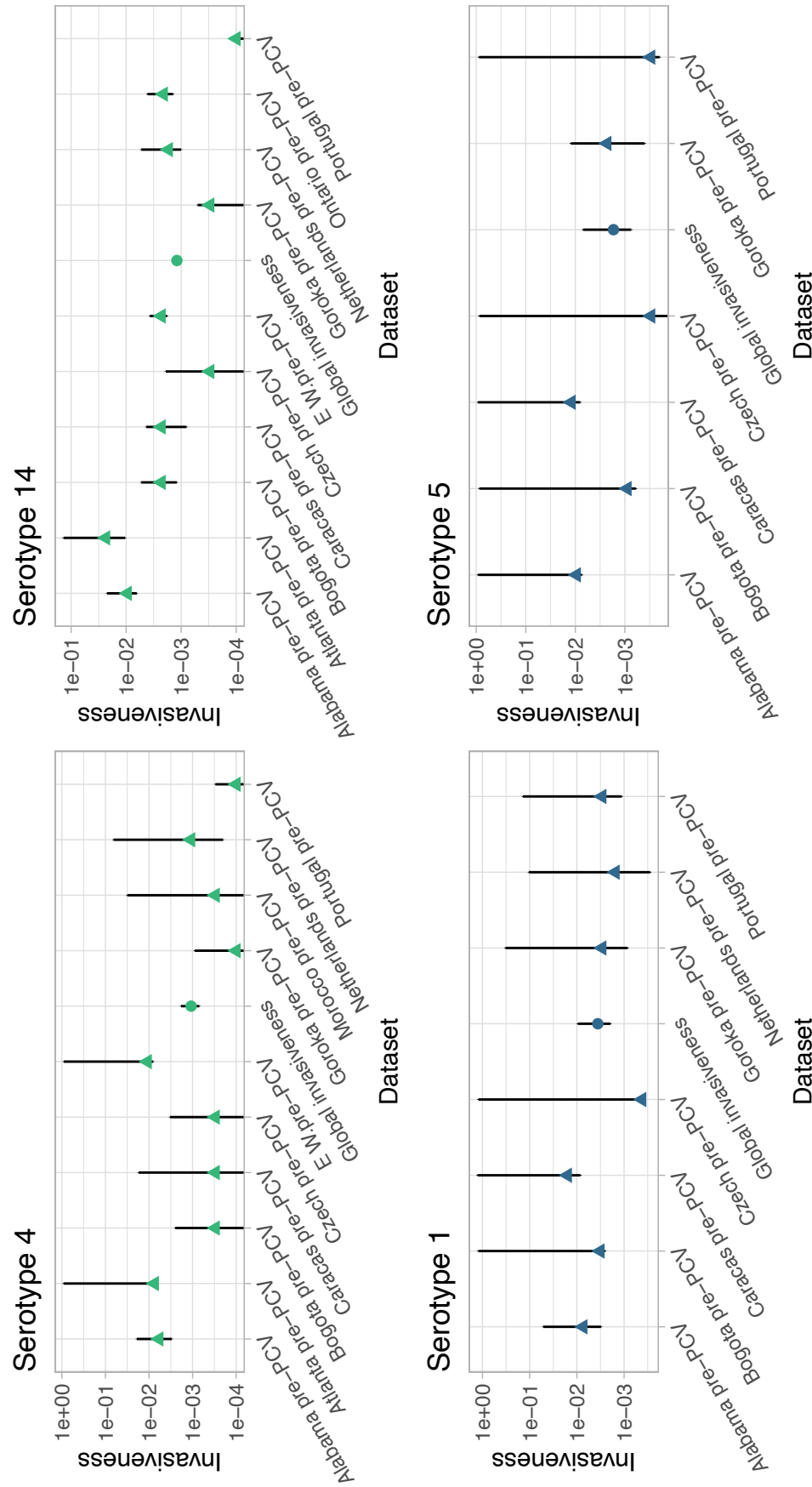
Figure 4.17: Invasiveness profiles of VT7 serotypes 4 and 14 and VT10 serotypes 1 and 5.
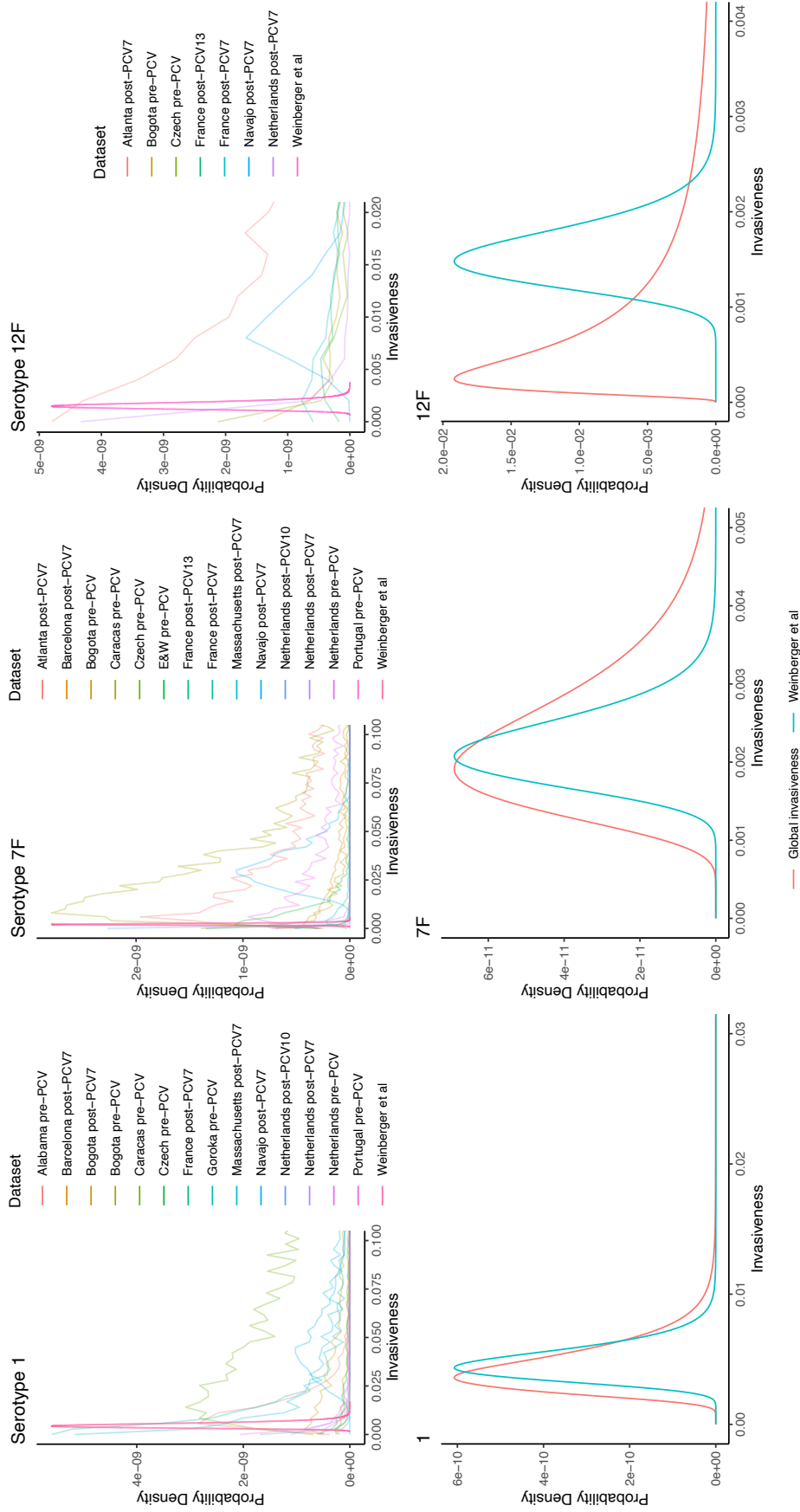
Figure 4.18: Comparison of posterior distributions of study-specific local invasiveness and global invasiveness from our models with those obtained by Weinberger et al.

4.4 Discussion

I developed a novel open source Bayesian framework for estimating both local and global invasiveness as the hazard of developing IPD given carriage of a particular serotype. The main finding was that substantial variation in pneumococcal invasiveness can be observed within a serotype, as shown by the variability between datasets (Figure 4.17), but the model comparison suggested that global invasiveness is favourable to local invasiveness when considering all datasets. The local invasiveness model would still be suitable for serotypes that are locally-emerging in the absence of global invasiveness estimates. The analysis found two other factors contributed to the hazard of IPD given carriage of a given serotype: host age, as some serotypes differed significantly in invasiveness between children and adults within the same setting and globally; and the location, as setting-specific differences in carriage and disease rates will directly impact invasiveness estimates.

In line with these results, previous studies have shown that invasiveness cannot be explained by serotype alone [355], though this is somewhat contested by studies estimating invasiveness through the odds ratio method (OR) [66]. Since the invasiveness estimates described in this chapter express an absolute rate that is independent of the other serotypes in a dataset, these estimates are not able to be directly compared with the OR method estimates. The OR method that does not use a standard serotype expresses a unitless measure for a dataset that is dependent on the other serotypes in that dataset, limiting the comparability to other datasets. Fitting an OR across datasets is undesirable given the variability of serotypes present between datasets. This is similar to the global invasiveness estimates presented here, which are estimated using variable combinations of datasets, making it difficult to compare the values between serotypes. Colijn et al. [331], whose meta-analysis was the basis for the datasets used in this chapter, fitted a random effects model across datasets to obtain logarithmic odds ratios on the basis of this method. Comparing the association between their fitted odds ratios and the global invasiveness estimates from this chapter indicates a positive and concordant relationship (Kendall's τ coefficient $= 0.61$, $p = 5.982 \times 10^{-8}$). This suggests some level of alignment on which serotypes are generally considered most invasive. The

second OR method relative to a standard serotype assumes that such a serotype will have the same invasiveness everywhere, which I have shown not to be the case for serotype 14. Brueggemann et al. estimated ORs relative to serotype 14 and found that the most invasive serotypes 1, 5 and 7 were around 60 times more invasive than the least invasive serotypes 3, 6A and 15 [66]. My analysis is in agreement with theirs on the highly invasive types, but serotypes 3, 6A and 15 (both 15A and 15B/C) were in the mid-range of invasive estimates relative to others. This is likely due to their characterisation of fewer serotypes. Additionally, contrary to Brueggemann et al.'s findings, serotype 8 was also identified as a highly invasive NVT in this analysis. Conclusions for serotype 8 may differ because of the relative nature of this OR method, which may not highlight invasive serotypes that would have similar invasiveness as the standard serotype.

Despite being preferred by BF when comparing all datasets together, the global invasiveness model may be somewhat limited in how well it would describe new datasets compared to the local invasiveness model. The more datasets included in the global invasiveness estimate, the narrower the posterior distribution and CrIs. This makes it less likely to describe corresponding serotypes in new and variable datasets, as seen with serotype 14. Additionally, the posterior distributions of global invasiveness estimates tend to be dominated by likelihoods that are more informative (e.g. narrower) rather than broader and diffuse [386]. As such, results suggest global estimates will provide a reasonable estimate of invasiveness in other settings in the absence of local invasiveness data, but local data and fits would be preferable to describe a specific setting. What this means is that this model, or definition of invasiveness, could either be missing a location-specific covariate, or that the data collection in the studies gathered by the meta-analysis was not consistent enough to be modelled accurately. This location-specific covariate would modify the global invasiveness model, such that all serotype invasiveness values in a setting would be moderated by setting characteristics such as healthcare availability, host genetics, and socioeconomic factors. This may reduce the sensitivity to varying local surveillance practices and potentially help us understand why locations vary in IPD burden, though its value would depend on the surveillance coverage

and accessibility of isolates. The challenges with such an estimate would be that it would require a breadth of data and serotypes, as well as a reference dataset that would make it a relative measure - something that is avoided with the method presented in this chapter. It would also not allow for a BF to be calculated per serotype. This analysis is out of scope for this thesis, but could be an interesting avenue for future research.

The variation in the posteriors for some serotypes also highlights setting-specific factors that may affect the invasiveness estimate. The heterogeneity in both local invasiveness and disease incidence between settings may have been attributable to location-specific factors such as lifestyle or healthcare disparities, including the prevalence of and prevention against comorbidities such as HIV [403], asthma, diabetes [404], cancer [405], influenza [406], among others. This emphasises an important limitation of this analysis, wherein it compared and combined studies from around the world that differ in socioeconomic development, access to healthcare, and other environmental factors. Differences in disease rates may also reflect contrasting rates of blood culture [407]. Rates of blood culture may depend on access to laboratories and healthcare centres, something that will vary between countries and even in rural versus urban settings. Additionally, the disease rates may have been driven by the data collection of different manifestations of IPD and different age categories, which we have consolidated here. Certain pneumococcal serotypes have been commonly associated with disease manifestations in specific age categories [64, 71, 366], which we grouped together here. Temporal biases may also have affected the local invasiveness. Fewer or more serotypes reported in the studies could reflect the year in which studies took place since more serotypes were discovered later and therefore more were distinguished in the serotyping procedures. Additionally, while studies were undertaken over different lengths of time, the results do not indicate a correlation between invasiveness and length of study and therefore the study period should not have a significant effect on invasiveness results. For example, the England and Wales study took place over one year and the Czech study took place over nine years, but they both had similar ranges of local invasiveness.

Invasiveness measures may also rely on other population-independent biological factors that

are still misunderstood, such as immune clearance and competition [62], and by extension
antibiotic resistance. For example, it has been suggested that increasing pathogen density
in the nasopharynx may promote bacterial invasion into the lungs, resulting in pneumonia
[408], and that viral coinfection can promote this process [409]. How specific serotypes may
promote this process is unknown. The local invasiveness model justified modelling the impact
of vaccination as being driven by changes in carriage, not by any change in invasiveness.
What this invasiveness measure may lack is the interaction between serotypes in a specific
environment, as all serotypes in the analysis were treated as independent. The question
would then be whether invasiveness is impacted by vaccination – something we were not
able to conclude given the limited post-vaccination data with no consistent trend between
datasets. For instance, it would be interesting to investigate whether the invasiveness of
serotype 19A has changed since vaccination in the same setting, considering it has been one
of the most prevalent serotypes in disease and has exhibited antibiotic resistance. The matter
is further complicated with a study showing that carriage alone does not fully explain the
variation in post-vaccination IPD [410]. A dynamic approach to invasiveness to explore how
it changes over time for each serotype could be one way of understanding the true impact of
vaccination or whether invasiveness is impacted by the presence of other serotypes.

An important limitation of this meta-analysis is the data on carriage. Contrary to IPD
incidence, the number of carriers is taken at distinct time points and does not reflect the
temporal trends of carriage loss and acquisition. This is because it is usually estimated as
the number of individual carriers over the study period divided by the total number of
individuals. Additionally, multiple serotype carriage is likely underestimated as multiple
carriage is difficult to detect with many traditional serotyping methods [2]. In this analysis,
multiple carriage was counted as multiple individuals, even though only one of the carried
serotypes is likely to be isolated from IPD. This means that some serotypes' invasiveness may
have been underestimated, since the probability of causing disease given carriage would be
lower if someone carried two serotypes but only one of these would usually cause disease
and be detected. Furthermore, variation in carriage between infants and older children

[67] will also affect the inferred carriage rates, which here aggregated children under 18 years. Serotypes will seem more invasive if carriage samples included children further from the peak age of carriage. Moreover, the assumption that adult carriage would mirror child carriage may be a limiting factor, as carriage prevalence has been shown to differ between the age groups [411]. Unfortunately, in the absence of widespread adult carriage data, this assumption was unavoidable.

It is unclear whether this study would help in vaccine design, had it been conducted prior to PCV7 manufacturing and licensing [331, 356, 412]. There is no evidence, from our limited pre- and post-vaccination analysis, that serotypes change invasiveness post-vaccination. Targeting the serotypes contributing to the most IPD first (i.e. 10 of 13 serotypes in the current PCV13) to alleviate the burden, followed by the highly invasive types (i.e. 8, 9N, 12F, 13, 18B, 20, 22F) to prevent replacement, as described in Chapter 3, may have theoretically reduced disease short- and long-term. Targeting the most invasive serotypes in children first (1, 4, 5, 7F, 8, 9V, 14, 18C, 19A, 22F, 27, 38) would also reduce the disease burden, particularly as many of these serotypes are also the ones that caused the most IPD globally [10]. However, this would likely also depend on the serotypes prevalent in carriage. For example, serotype 3 which had a higher invasiveness in adults than children (although non-significant), still causes considerable disease in adults despite its inclusion in PCV13 [163, 413, 414]. Previous studies have shown that based on universal invasiveness ratios, current vaccine design is not optimal as it depends heavily on serotypes in circulation, which differ by population, and which age groups are targeted [331, 356, 412]. Current vaccine design may be more problematic because of the heterogeneity seen in the posteriors of the local invasiveness estimates. Regardless, this could be explored in a dynamic transmission model that could simulate the better vaccine, as well as to understand our current situation: given the implementation of PCV7, PCV10 and PCV13, and additional vaccines being developed that build on those formulations, what will be the difference between adding serotypes with high IPD incidence versus adding those with a high local invasiveness? This would be interesting to explore for different locations and age groups given the variability between locations.

Ultimately, a novel methodology has been employed to estimate invasiveness using a multitude of datasets. As the code and data are publicly available [357], this analysis is expandable to other datasets, allowing others to estimate local invasiveness estimates for their own datasets and to re-estimate the global invasiveness model. Additional studies on cases of carriage and disease in a population pre- and post-vaccination in both age groups could make these results more robust and may also shed light on new or rare invasive serotypes post-vaccination. This method could be used to gather data in the future and allows the invasiveness of newly emerging serotypes to be determined. In line with previous results, these results indicate that while PCVs target serotypes and serotypes are not the sole contributor to influencing invasiveness [355], PCVs targeting serotypes may not be effective in alleviating the disease burden long-term globally [331]. Since serotype is not the only determinant of invasiveness, protection against invasive strains may not be conferred globally from PCVs of expanded valency. As the model agrees that there is substantial variation in invasiveness across the species, it is unclear whether it would be preferable to target serotypes with high local invasiveness with setting-specific vaccines or to target those with high global invasiveness worldwide in future PCV design. In essence, this strengthens the evidence for a vaccine targeting invasive strains and not serotypes alone.

4.5   Aim in this thesis

Using a Bayesian approach, I estimated a global invasiveness rate for each serotype across all datasets as well as a local invasiveness. Serotypes with the highest global invasiveness were 5, 1, 4 and 7F in children, 8, 4, 1 and 20 in adults pre-vaccination, and 1, 7F and 12F in children post-vaccination. There was no correlation between the global invasiveness pre-PCV and the number of GPSCs in which a serotype was found in children, but a negative correlation in adults. Local invasiveness was inconsistent across studies, something that may be attributable to population heterogeneity, and suggests that invasiveness may not be universal. Additionally, the most invasive serotypes varied between datasets, potentially as a result of biological differences such as competition and interaction between serotypes.

Model comparison showed that the global invasiveness model was preferred to the local

invasiveness model. However, this is likely due to the reduction in the number of parameters

in the model, as the global invasiveness model does not always capture the variability

between datasets. Going forward, invasiveness estimates could investigate and incorporate

population features that give rise to this variability in order to be more useful for vaccine

policy. In the next chapter, the local and global invasiveness estimates from this chapter will

be used to understand the impact of targeting serotypes based on their invasiveness, both

local and global, versus targeting serotypes based on their disease incidence.

Chapter 5

Impact of expanding PCVs based on different vaccine strategies: an individual-based dynamic transmission modelling study

This chapter evaluates the impact of increasing PCV valency in European and North American settings using an individual-based dynamic transmission model. The three strategies of vaccine expansion are disease incidence-based, local invasiveness-based and global invasiveness-based.

Dissemination

A modified version of this chapter is in preparation as

Løchen A, Truscott JT, Croucher NJ, Anderson RM. An individual-based dynamic transmission model to determine the impact of increasing serotype coverage in pneumococcal conjugate vaccines.

5.1 Introduction

While many transmission models have been developed to evaluate the impact of implement-
ing PCV (Chapter 2), none have used a serotype-specific model to estimate the impact of
increasing the serotype coverage of PCV according to different serotype targeting strategies.
A previous study estimated the optimal serotype composition in the Finnish context, although
it did not account for transmission dynamics or serotype-specific durations of carriage which
would likely affect how serotypes behaved post-PCV [356]. These dynamics are important
to explore as PCVs are expanded globally to address serotype replacement. This chapter
aims to evaluate this impact of different serotype-targeting strategies by implementing an
individual-based model that takes into account serotype-specific and non-specific immu-
nity, to reproduce the stable coexistence of multiple serotypes, and vaccination with PCVs
of increasing valency. As observed in Chapter 3, there is evidence for differences in post-
vaccination serotype landscape in high-income countries that may impact an implemented
vaccine's effectiveness. The currently licensed vaccines were based on serotypes causing the
most disease in the US (VT7) and expanded based on epidemic serotypes causing disease in
LMICs (VT10), post-vaccination serotype replacement (VT13), or antibiotic resistance (VT13).
In this chapter I wanted to understand which serotypes would be most beneficial to add in
future PCVs, whether based on local invasiveness, global invasiveness, or disease incidence,
and the effect across countries in North America and Europe.

5.2 Methods

5.2.1 Data and parameter fitting

To construct the model, pre-vaccination carriage data and previously obtained invasiveness
estimates (Chapter 4) were used to estimate the impact of increasing serotype valency in
France and the United States of America (USA), chosen as representatives of Europe and

North America. Country-specific social contact matrices were used [415, 416], and the population was split into children (< 5 years), adults (5-64 years), and elderly (65+ years) as per the contact matrices.

Previously described pre-vaccination carriage data from an infant (< 5 years) refugee population in Southeast Asia [119, 417–419] was used to fit carriage duration and effective contact rate parameters. Briefly, monthly swabs were taken from a Maela longitudinal birth cohort over the course of twenty-four months using both latex sweeping and the World Health Organization (WHO) method, resulting in over 4500 total carriage isolates and a total of 69 serotypes modelled. From this monthly swab data, the duration of carriage episodes accounting for censoring and swabbing resolution was taken from a previously published study [419]. Non-typeable pneumococci (NT) were also included in the model, as they are prevalent in carriage, although they do not typically cause invasive disease.

While carriage duration was already estimated from this data by Lees et al [419], the authors used a Hidden Markov Model (HMM) which did not depend on an individual's serotype-specific previous colonisations, something which was considered vital to building immunity in the model described here. Duration of carriage of serotype i, ϕ_i, were therefore assumed to follow a gamma distribution with subsequent colonisations of that serotype:

$$\phi_i = \text{Gamma}(\alpha_i, \beta_i) \tag{5.1}$$

$$\beta_i = \frac{\alpha_i}{\mu_i} \tag{5.2}$$

$$\mu_i = A_i \exp(-\psi_i \times B_i) \tag{5.3}$$

where α_i is the shape parameter of the gamma distribution, μ_i is the mean of the gamma distribution, and ψ_i is the number of previous colonisations with that serotype. α_i, A_i and B_i parameters were estimated using the Python MCMC sampler package emcee [420] with uninformative priors. Recovery rates are the inverse of the mean carriage durations.

The density-dependent transmission coefficient, τ is equivalent to the initial force of infection
(FOI) divided by the number of infectious people (τ=FOI/I) [421], which was fixed to an
average 40% prevalence, as in the Cobey and Lipsitch model [57]. Initial FOI was estimated
from the data as the number of carriage acquisitions in the data population divided by the
product of the total data population and length of the carriage study.

For model validation, the maximum likelihood estimates of the carriage duration parameters
were inserted into a simulation of the data (i.e. naïve, or fully susceptible) population to
check the overall carriage prevalence was within 1% of the target carriage prevalence (40%).
Additionally, the distribution of previous colonisations in the data population was compared
to that of the simulated population for the same time period (twenty-four months) using the
fitted epidemiological parameters.

To obtain serotype-specific colonisation history distributions, a model with a naïve child (< 5
years old) population was simulated to run for 90 years. After the initial 20 years, the model
recorded the proportion of each number of previous colonisations in the population every
five years to obtain a distribution of previous colonisations across five-year age bands. This
was run multiple times and averaged to obtain the final serotype-specific distributions of
previous colonisations. Children between 5 years and 17 years were initialised to not have
had previous colonisations. Model parameters are found in Table 5.1.

5.2.2   Model

An individual-based model (IBM) was chosen to avoid the increasing number of compart-
ments and combinations inherent to multi-strain multi-carriage compartmental models. IBMs
are also useful for describing within- and between-host interactions [422]. The full model
is described in detail as per the Overview, Design concepts and Details (ODD) standard
protocol [423] (Appendix B) and is represented diagrammatically in Figure 5.1.

A number of key assumptions were made in the model regarding immunity, competition

and vaccination, many in line with the Cobey et al. study [57]. To summarise, specific immunity was assumed to depend on the number of previous serotype-specific colonisations and therefore impacted the serotype-specific FOI felt by an individual. Non-specific immunity adjusted serotype recovery rates in conjunction with the total number of previous colonisations (regardless of serotype). Competition was fixed across serotypes (i.e. serotype-independent) and altered an individual's serotype FOIs only if they were already carrying a serotype (but were below the maximum carrying capacity). Once the maximum carrying capacity was reached, an individual could no longer acquire serotypes. Vaccines were assumed to be all-or-nothing, i.e. the vaccine efficacy informed the proportion of individuals vaccinated that would be successfully immunised and the remainder would be unprotected. Vaccination did not feature waning protection over time, but protection ceased altogether after the specified duration of protection.

The model was programmed in Python and run on a range of Intel Xeon E5-series machines with between 12 and 24 cores, and 96-256Gb of RAM, running Windows Server 2012 R2, as part of a Microsoft High Performance Computing (HPC) 2012 R2 cluster. Post-simulation analysis was undertaken in R.
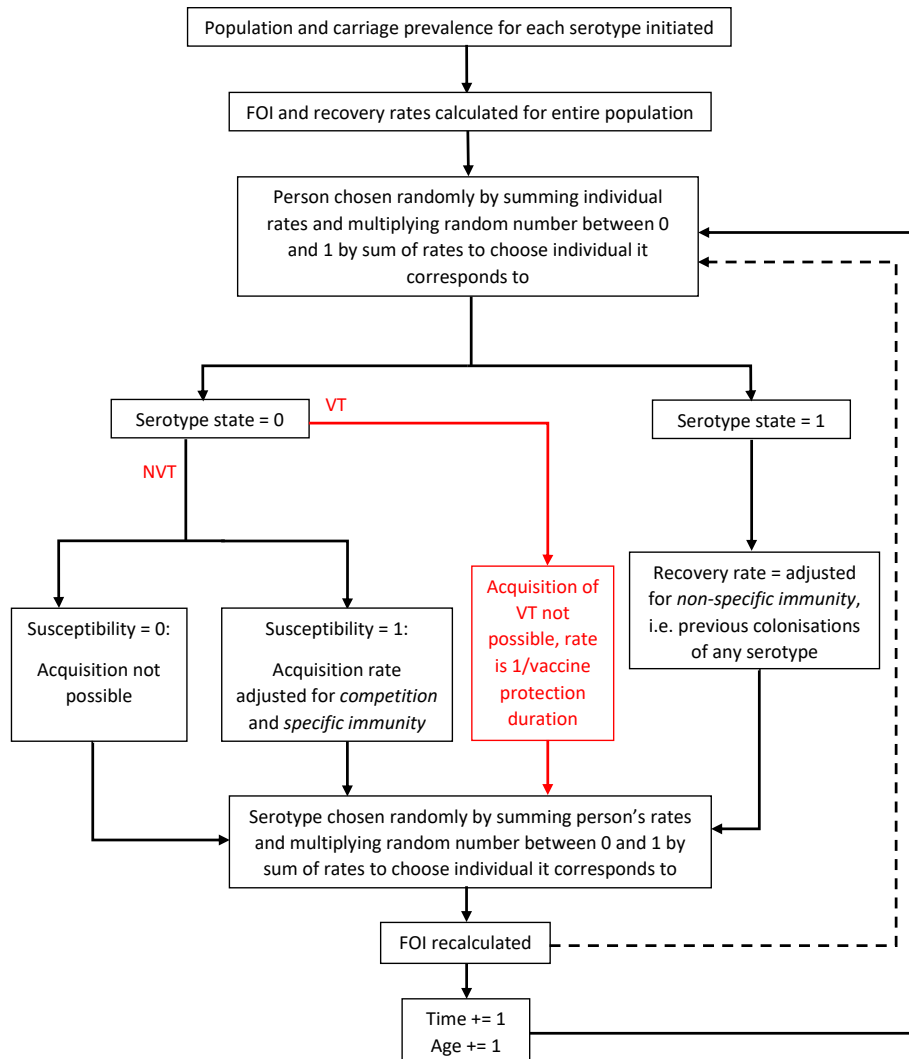
Figure 5.1: Diagram of individual-based stochastic model of pneumococcal transmission and vaccine impact. In red, vaccination prevents acquisition of VT and therefore the rate is the rate of loss of vaccine-induced protection. VT: vaccine serotypes; NVT: non-vaccine serotypes; FOI: force of infection. Dashed line represents multiple switching events each time step.

Briefly, a population of 5000 individuals with an initial carriage prevalence of 1% for all serotypes was simulated. As previously mentioned, the population was structured into three age groups: children less than five years, "adults" between five years and 64, and the elderly aged 65 years and over. Population breakdowns by age group were obtained from the United Nations [395], with the French population (and American) broken down as 6% children under five (6%), 77% adults (81%), and 17% elderly (13%). The population

was closed – i.e. the population number stayed constant and each time a person reached the average life expectancy of a setting [424], they died and were replaced with a newborn [211, 425]. This was assumed to balance out the population of children and elderly, as even though children spend a shorter amount of time in their age category (5 years) compared to the elderly (13 to 15 years), the populations of children less than 5 years in France and the US are less than half that of the elderly. Each person was either susceptible to (0), or carrying (1), each serotype, but could carry no more than two serotypes at a time. This maximum carrying capacity, m, was chosen to avoid too much complexity in the numerical simulations (and hence computational burden), while also adhering to the infant refugee camp data which showed that co-colonisation occurred in at least 10% of cases [81]. This is also the most common number of serotypes detected in multi-serotype carriage [426]. Each person could, depending on whether they were a carrier or not, either recover from carriage ($1 \rightarrow 0$) at the recovery rate or get infected with the serotype ($0 \rightarrow 1$) at the FOI. For each serotype i and for individuals in age group k, $FOI_{i,k}$ is defined as the product of the per contact transmission probability τ (i.e. infections per contact), the average number of contacts per time between individuals in age groups j and individuals in age group k, $c_{j,k}$, and the number of people infected with serotype i in age group j, $I_{i,j}$ divided by the population in age group j, N_j:

$$FOI_{i,k} = \tau \sum_j c_{j,k} \frac{I_{i,j}}{N_j} \tag{5.4}$$

The model worked as follows: first, the FOI was calculated for each serotype based on the number of people infected with that serotype. For each serotype, each individual was assigned either the FOI rate, if they were susceptible (0), or the recovery rate, if they were carriers (1), like so:

	Serotype A	Serotype B
Person 1	0	0
Person 2	1	1

Carriage data frame

$\longrightarrow$

	Serotype A	Serotype B
Person 1	FOI_A	FOI_B
Person 2	γ_A	γ_B

Rates data frame
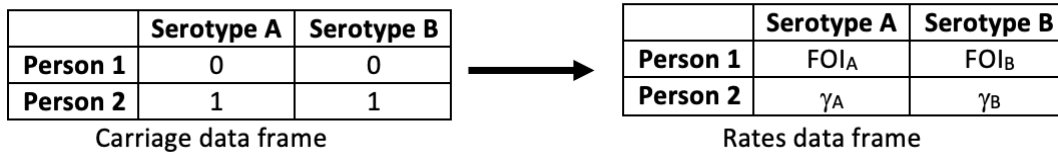
Figure 5.2: Representative example of process of serotype rate assignments. Person 1 is carrying neither serotype, and therefore both serotypes are assigned their respective forces of infection (FOI). Person 2 is carrying both serotypes, and therefore both serotypes are assigned their respective recovery rates, γ_i.

Choosing the individual to which a serotype switching event happens

Each time step was set at one day, and there were $N/50$ serotype switching events per time step, in which an individual's carriage status changes from $0 \to 1$ for acquisition or $1 \to 0$ for recovery. The optimal number of events per time step was obtained during the naïve child population simulations through trial and error: originally the model was set up with asynchronous time where the natural logarithm of a random number drawn from a uniform distribution was divided by the sum of all rates. Because this extended the simulation running time significantly, the number of events per time step were approximated from these simulations as a function of the population. To choose the person affected by a serotype switching event, a random number was drawn from a uniform distribution $x \in U(0,1)$ and multiplied by the sum of all individuals' rates Σ:

	Serotype A	Serotype B	Sum
Person 1	FOI_A	FOI_B	$FOI_A + FOI_B = \Sigma_1$
Person 2	γ_A	γ_B	$\gamma_A + \gamma_B = \Sigma_2$
			$\Sigma_1 + \Sigma_2 = \Sigma$

Figure 5.3: Representative example of process for summing individuals' rates before choosing the individual that the serotype switching event will happen to.

When compared to the cumulative sum of individuals' rates Σ, the chosen person corresponded to the rate where the number fell in the interval. In the example above, this would correspond to Person 1 if $x * \Sigma$ fell within the interval 0 to Σ_1 and would correspond to Person 2 if $x * \Sigma$ was greater than Σ_1 and less than Σ_2.

Altering an individual's serotype rates with colonisation history and choosing the serotype to which the switching event happens

The chosen person would then have their serotype rates altered depending on how many serotypes they were currently carrying and their colonisation history. If they were carrying the maximum carrying capacity, then all other serotype FOI rates became 0 such that no additional serotype could be acquired. Otherwise, the FOI of serotype i for the chosen person changed as follows [57]:

$$FOI_i' = \begin{cases} 0 & \text{if } s = 0 \\ FOI_i & \text{if } \psi_i = 0 \text{ and } s = m \\ (1 - \sigma)FOI_i & \text{if } \psi_i > 0 \text{ and } s = m \\ (1 - \theta)FOI_i & \text{if } \psi_i = 0 \text{ and } 0 < s < m \\ (1 - \sigma)(1 - \theta)FOI_i & \text{if } \psi_i > 0 \text{ and } 0 < s < m \end{cases} \tag{5.5}$$

where σ is the serotype-specific immunity that is applied if the number of previous colonisations of serotype i, ψ_i, is greater than 0; θ is the competition parameter that is applied if the susceptibility score s, i.e. how many more serotypes can be carried, is less than the maximum carrying capacity m. In other words, if $s = m$, m additional serotypes can be carried but no competition parameter is applied because there is no resident serotype colonising the individual. If $s = 0$, zero additional serotypes can be carried and therefore the FOI of all non-colonising serotypes becomes 0. Previous studies have shown that the mechanism of reduction in multiple carriage from between-serotype competition is through a reduction in acquisition [288, 427], therefore the competition parameter reduces the FOI if an individual is already carrying. This is done in a serotype-independent manner, i.e. the competition parameter is assumed to be constant across serotypes.

The equation for the recovery rates of serotype i being carried by an individual was obtained from the Cobey and Lipsitch study [57]:

$$\gamma_i = \frac{1}{\kappa + [\phi_i - \kappa]\exp\left(-\varepsilon \sum_i \psi_i\right)} \tag{5.6}$$

where κ is the minimum duration of carriage, which was set to 5 days per the longitudinal carriage data; ϕ_i is the mean duration of carriage of serotype i in the absence of immunity, ε is the non-specific immunity parameter; and $\sum_i \psi_i$ is the sum of the individual's number of previous colonisations (of all serotypes). These serotype-specific recovery rates serve as a proxy for serotype fitness, as their transmission rates are equivalent and therefore their carriage durations determine their prevalence [57]. Parameters are found in Table 5.1.

The serotype to which the switching event happened was chosen using the same method as the person-choosing event. If a serotype was acquired in the serotype-choosing process, the number of previous colonisations was updated for the individual, unless they were in the elderly age category as it was assumed that individuals lose immune memory after the cut-off age.

Parameters to maintain serotype diversity

Because serotypes die out naturally due to the nature of stochastic individual-based models and the dynamics of multi-strain pathogens, the model was re-seeded with 1 child carrier each time a serotype became extinct to maintain the low levels in carriage of many serotypes due to migration, similar to a previous study [254].

There is evidence for two mechanisms by which immunity to pneumococcal infection is generated: anti-capsular (serotype-specific) immunity and non-specific [46, 56, 57, 254]. Serotype-specific immunity reduces the likelihood of re-acquisition of a certain serotype, enhancing serotype competition and potentially providing a mechanism to maintain serotype diversity (Equation 5.5). In this model, it is a reduction in the FOI in subsequent colonisations for the serotypes that were already carried by the host. Non-specific immunity is evidenced by the higher carriage prevalence in children, particularly infants [56]. In this model, it is a

Table 5.1: Summary of model parameters.

Parameter	Value	Description	Reference
$c_{j,k}$	-	Average number of contacts age group j has with age group k	[416] for USA (UK data used), [415] for France
ϕ_i	-	Fitted carriage duration of serotype (days) i	Data from [419]
τ	-	Risk of transmission (per contact), fitted to 40% carriage prevalence	Cobey & Lipsitch [57]
σ	0.4	Specific immunity	Cobey & Lipsitch [57]
ϵ	0.1	Non-specific immunity	Cobey & Lipsitch [57]
θ	0.01	Competition	Cobey & Lipsitch [57]
κ	5	Minimum carriage duration (days)	Data from [419]
v	95%	Vaccine efficacy	[154]
ω	10 years	Mean duration of vaccine protection	[429]

Table 5.2: Stochastic event table showing changes in the number of susceptible (ΔS_i) and infected (ΔI_i) individuals over one switching event with serotype i for a person in age group j.

Event	Change ($\Delta S_i, \Delta I_i$)	Probability per switching event
Infection	(-1, +1)	$\tau \sum_j c_{j,k} \frac{I_{i,j}}{N_j}$
Recovery	(0, -1)	$\gamma_i(\kappa, \phi_i, \epsilon, \psi_i)$

reduction in the duration of carriage, i.e., a faster recovery rate. Elderly individuals in the model do not have their number of previous colonisations updated after acquisition to reflect the impaired immunity with age [428]. Because serotypes have fitness differences in their recovery rates (hence differing basic reproductive numbers R_0), the model described here does not meet the criteria for being a neutral null model [290]. Co-existence is maintained through the migration rate and these two mechanisms of immunity that enhance diversity.

The stochastic events for serotype i are summarised in Table 5.2.

Simulations were run 20 times each. At the end of a simulation, disease incidence was estimated from local invasiveness estimates from the respective countries' datasets (Figure 5.4), assuming that invasiveness did not change post-vaccination. The invasiveness datasets used were Massachusetts post-PCV7 for the USA [382] and France post-PCV7 for France [381].

The Massachusetts invasiveness dataset describes serotype 6A/C, which here was considered equivalent to serotype 6A. Where the country's invasiveness dataset did not contain a serotype included in the simulation, the global invasiveness was taken instead. In the case where the serotype was not included in neither local nor global invasiveness datasets, a low invasiveness of 0.0001 cases of invasive pneumococcal disease (IPD) per 100,000 carriers per year was assigned.
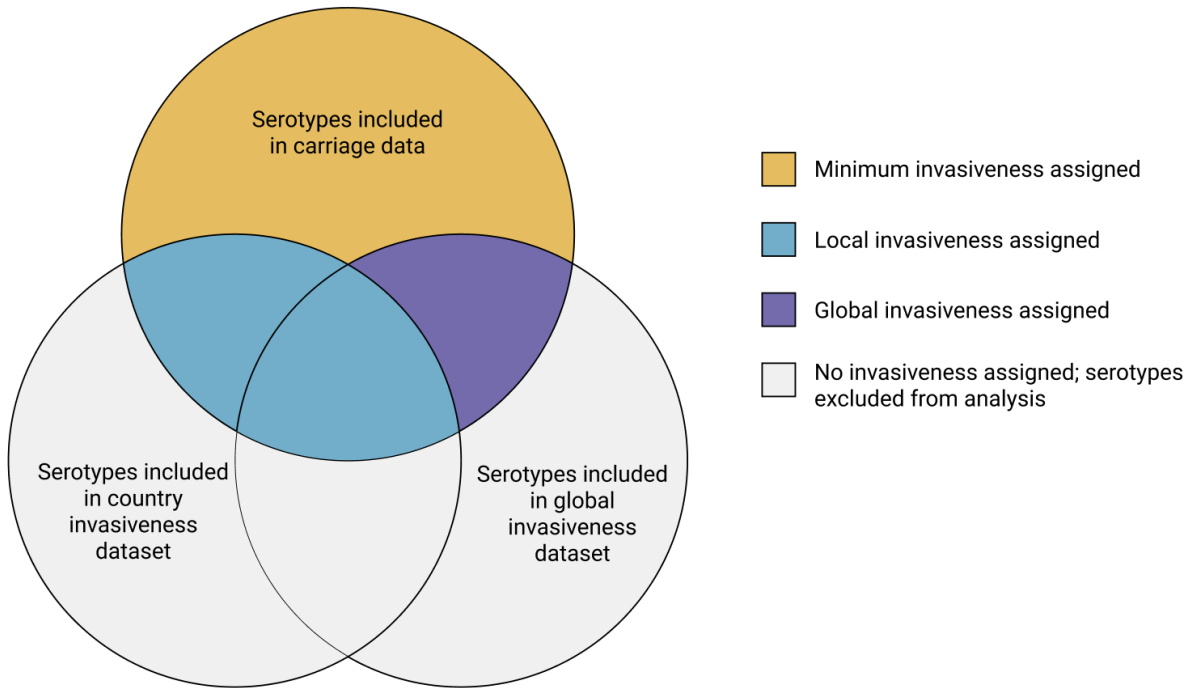
Figure 5.4: Illustration of serotype data and corresponding invasiveness according to availability. Country-specific invasiveness was used where available. In its absence, global invasiveness was used. If both local and global invasiveness values were missing, a low invasiveness of 0.0001 cases of IPD per 100,000 carriers per year was assigned.

5.2.3   Vaccination simulations

Vaccination with PCV usually occurs from one year of age with multiple doses up to two years of age. To simplify the dosing schedule and account for the time for immune responses to build, vaccination was implemented on a rolling basis by protecting children aged between two and three years after 20 years in the simulation, with a 95% vaccine efficacy (v) against all serotypes included in the vaccine [154] and protection lasting a duration of ω, 10 years

[429]. Vaccine-induced protection was considered to be 'all or nothing', i.e. there was no decay in the protection modelled and successfully immunised individuals (v, or 95% of those vaccinated) were considered to be fully protected until the duration of protection. In other words, $(1 - v)$ proportion of the vaccinated individuals were not protected at all. After a single pre-vaccination period, three vaccination strategies were investigated:

1. Disease incidence strategy: The serotypes most prevalent in causing disease were included in the vaccine. This was calculated by applying local invasiveness estimates (and global invasiveness estimates in their absence) to the carriage prevalence to obtain serotype-specific disease incidence. This strategy was only implemented in the French setting.

2. Local invasiveness strategy: The serotypes with the highest local invasiveness, based on the French invasiveness values, were included in the vaccine. This would target the serotypes that have the highest hazard of causing disease given carriage in a specific location. This strategy was only implemented in the French setting to compare the impact of targeting locally invasive serotypes rather than highly prevalent serotypes.

3. Global invasiveness strategy: The serotypes with the highest global invasiveness values were included in the vaccine. This would target the serotypes that have the highest hazard of causing disease given carriage across global settings in Chapter 4. This strategy was implemented in both the French and American settings to compare the differences in carriage dynamics and disease incidence post-vaccination with the same formulations.

Having a common pre-vaccination period allowed for a better comparison between the strategies within a country. As in reality, vaccination was not considered to be effective against serotypes already being carried [430]. In the first instance, each scenario was run for 10 years with thirteen serotypes included in the vaccine (VT13), corresponding to a hypothetical PCV13 (hPCV13). After this, a higher valent vaccine was introduced that covered an additional seven serotypes (VT20) corresponding to hypothetical PCV20, hPCV20,

chosen in the same fashion as the first hPCV in the respective scenarios. This was then run for 10 years before the introduction of the following higher valent vaccine (hPCV30) which covered ten additional serotypes (VT30).

At the end of a simulation, the incidence rate ratios (IRRs) were estimated for each age group using the IPD incidence before and after vaccination with each individual hPCV as:

$$IRR_{hPCV} = \frac{IPD_{post-hPCV}}{IPD_{pre-hPCV}} \tag{5.7}$$

and used the standard deviation based on the number of disease cases d before and after vaccination to estimate the 95% confidence interval (CI) [431, 432]:

$$\overline{SD}[\ln(IRR_{hPCV}] = \left(\frac{1}{d_{post-hPCV}} + \frac{1}{d_{pre-hPCV}}\right)^{1/2} \tag{5.8}$$

$$\underline{IRR_{hPCV}}, \overline{IRR_{hPCV}} = \exp(\ln(IRR_{hPCV}) \pm 1.96 \times \overline{SD}[\ln(IRR_{hPCV})]) \tag{5.9}$$

where $\underline{IRR_{hPCV}}$ and $\overline{IRR_{hPCV}}$ indicate the lower and upper bound of the 95% CI, respectively. These values were also estimated for the overall simulation period, comparing the pre-hPCV period and the post-hPCV30 period for each age group. If zero cases of disease were detected in an age group, a correction of 10^{-8} was added to d to avoid division by zero errors in the standard deviation calculation [433]. These age-specific IRRs were then fitted to a fixed effects model using the meta package in R [434] to estimate the effect size of each model over the simulations and the respective 95% CI. All code is available on Github [435].

5.2.4   Sensitivity analysis

Due to the uncertainty in the values of the parameters listed, a scenario-based approach was taken for most parameters (Table 5.3). Because of the computational complexity of the model, populations of 500, 1000 and 5000 individuals were also tested to evaluate the stability of

Table 5.3: Alternative parameters used for scenario-based model analysis.

Parameter	Default value	Alternative values tested	Description	Reference
N	5000	500, 1000, 10,000	Total individuals in population	
θ	0.1	0.2, 0.5	Competition	[220, 235, 243, 436]
ε	0.1	0.25, 0.4	Non-specific immunity	[57]
σ	0.4	0.6, 0.8	Serotype-specific immunity	[57]
v	95%	75%, 85%	Vaccine efficacy, considered 'all or nothing' (Chapter 2)	[154, 164, 429, 437–439]
ω	10 years	5 years, 25 years	Vaccine duration of protection	[237, 251, 429]

the model, such that a smaller population could be reliably employed in the simulations to obtain general insights by comparison with the numerical results generated in a larger 10,000 population.

5.2.5 Comparison with another setting

The global invasiveness strategy was also implemented in an American setting for comparison with the French setting, with both the default parameter set and the sensitivity scenarios, to compare differences in carriage dynamics and post-vaccination disease incidence. Population-level differences were primarily in the disease incidence as a result of differing local invasiveness values, since maximum life expectancy and contact matrices do not differ drastically between France and the US.

5.3 Results

5.3.1 Estimation of epidemiological parameters

MCMC parameter estimates for serotype recovery rates are found in Table 5.4, along with
maximum carriage durations. Examples of the parameter distributions are also shown for
serotype 4 and serotype 19F (Figure 5.5, Figure 5.6). The transmission probability per contact
τ for a population with an average 40% carriage prevalence was equivalent to 1.21×10^{-4}.
The distribution of previous colonisations largely matched the previous colonisations from
the model when testing these epidemiological parameters in a naïve population (Figure 5.7),
and the resulting overall carriage prevalence was 39.9%. Unsurprisingly, the data have a
larger variance relative to the modelled distribution because the other serotypes can influence
parameters in a way that the inference done will not capture since the inference of recovery
rates was independent of other serotypes. In essence, the fitting does not account for the
interdependence of values in the full model from parameters such as non-specific immunity
and competition.

Table 5.4: Markov chain Monte Carlo parameter estimates and the resulting maximum
carriage duration for each serotype from a fitted gamma distribution. The gamma distribution
is parameterised with the shape parameter α and the rate parameter β. β is a function of
α, the A and B parameters, and the previous number of serotype-specific colonisations ψ_i
(Equation 5.2, Equation 5.3).

	A		B		α		Max carriage duration (days)
	Mean	Stdev	Mean	Stdev	Mean	Stdev	
4	105.2	12.1	0.0	0.2	1.9	0.4	86.8
6B	136.0	9.1	0.0	0.0	1.7	0.1	131.3
9V	115.2	33.0	0.0	0.1	1.7	0.4	45.9
14	115.3	21.5	0.0	0.1	2.2	0.3	60.0
18C	109.8	37.3	0.0	0.3	1.9	0.5	98.0
19F	206.5	12.2	0.0	0.0	1.5	0.1	213.9
23F	154.3	11.1	0.0	0.3	1.4	0.2	321.7
1	109.7	11.4	0.1	0.1	1.9	0.6	37.6
5	108.1	14.0	0.0	0.2	1.4	0.4	73.1

	A		B		α		Max carriage duration (days)
	Mean	Stdev	Mean	Stdev	Mean	Stdev	
7F	115.0	16.2	0.0	0.3	1.8	0.6	71.5
3	106.5	16.6	0.0	0.3	1.6	0.4	194.6
6A/C	106.8	5.3	0.0	0.2	1.5	0.3	28.9
19A	108.5	18.3	0.0	0.1	1.9	0.3	72.8
10A	115.1	17.0	0.0	0.2	2.0	0.5	104.9
10B	114.6	11.5	0.0	0.2	2.0	0.4	44.7
10F	122.3	20.6	0.0	0.2	1.8	0.5	37.6
11A	103.2	5.8	0.0	0.2	1.9	0.4	94.7
12B	201.6	57.6	0.3	0.3	2.5	1.4	298.6
12F	125.7	34.8	0.0	0.1	1.9	0.9	135.1
13	106.8	18.8	0.0	0.1	1.9	0.4	14.2
15A	105.7	14.9	0.0	0.1	2.2	0.4	110.4
15B/C	199.9	58.1	0.0	1.2	2.4	1.4	481.7
16F	117.5	15.0	0.0	0.3	1.8	0.5	82.7
17F	110.8	10.3	0.0	0.3	2.2	0.7	69.1
18A	119.3	23.4	0.0	0.2	2.2	0.9	39.9
18B	200.6	57.5	0.1	0.1	2.5	1.4	75.2
18F	154.8	51.0	0.1	0.0	1.8	1.1	102.1
19B	105.5	6.3	0.0	0.0	1.5	0.3	212.5
2	200.5	57.6	0.5	0.4	2.5	1.4	237.5
20	110.2	12.4	0.0	0.3	1.5	0.4	74.6
21	113.0	13.4	0.0	0.1	2.0	0.4	82.3
22A	115.5	17.5	0.0	0.2	2.3	0.8	293.5
22F	111.2	17.1	0.0	0.2	2.9	0.9	87.4
23A	105.1	5.9	0.0	0.2	1.6	0.3	100.6
23B	114.8	12.7	0.0	0.4	1.6	0.4	96.8
24A	182.7	57.1	0.2	0.2	2.0	1.3	102.9
24F	108.3	13.0	0.0	0.2	1.8	0.5	13.7
25A	177.3	57.4	0.0	0.1	1.7	1.2	72.7
25F	192.5	57.9	0.0	0.0	2.7	1.3	140.4
28A	183.1	54.6	0.6	0.3	2.6	1.2	212.0
28F	106.7	10.3	0.0	0.3	2.0	0.6	16.5
29	140.7	40.6	0.1	0.1	2.6	1.0	234.0

|	A		B		α		Max carriage duration (days)
	Mean	Stdev	Mean	Stdev	Mean	Stdev	
31	114.3	17.5	0.1	0.0	2.0	0.7	6.0
32A	131.1	34.5	0.3	0.1	2.2	0.9	78.8
32F	173.4	55.8	0.2	0.2	2.0	1.1	447.7
33B	118.6	14.5	0.0	0.2	1.9	0.5	104.8
33C	122.1	27.1	0.0	0.3	1.8	0.6	23.9
33D	137.2	38.4	0.0	0.1	1.7	0.6	208.7
33F	119.6	23.6	0.0	0.2	3.2	1.0	105.7
34	108.9	28.3	0.0	0.1	1.6	0.3	38.2
35B	200.3	58.0	0.2	0.1	3.0	1.3	161.0
35C	105.6	13.1	0.0	0.2	1.6	0.3	330.0
35F	110.0	10.9	0.0	0.4	1.5	0.4	11.2
36	201.0	57.5	0.1	0.1	2.6	1.4	43.7
37	146.4	44.9	0.1	0.0	2.0	0.9	156.6
38	104.8	14.6	0.0	0.2	2.9	0.7	93.1
39	157.6	49.9	0.1	0.0	2.3	1.2	49.6
40	115.8	20.0	0.0	0.1	2.0	0.8	163.8
41F	164.6	53.3	0.3	0.2	2.7	1.3	108.4
42	198.7	57.8	0.5	0.4	2.6	1.4	337.6
45	121.7	20.2	0.1	0.3	3.7	1.0	130.9
46	107.7	9.2	0.0	0.3	2.1	0.7	82.9
48	193.1	54.6	0.1	0.2	3.7	1.1	125.6
7B	115.0	19.2	0.0	0.2	1.9	0.6	264.3
8	112.5	25.7	0.0	0.4	1.7	0.5	60.5
9A	213.4	51.9	0.1	0.0	1.9	0.9	29.4
9L	127.4	31.0	0.0	0.3	1.9	0.8	152.1
9N	108.7	18.4	0.0	0.2	2.6	0.6	67.7
NT	136.3	11.7	0.0	0.3	1.1	0.2	88.1
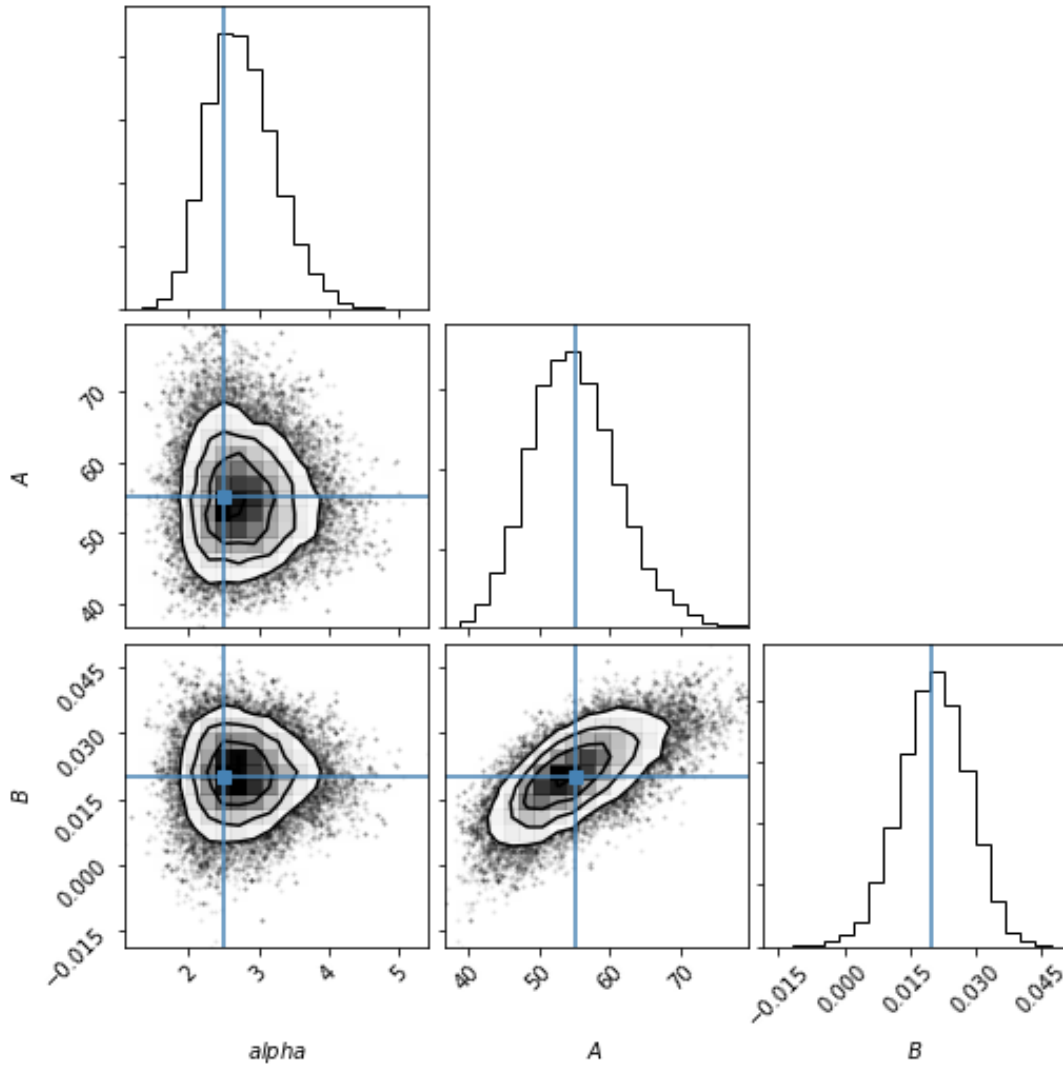
Figure 5.5: Scatter plot matrix of parameter distributions for serotype 4. Diagonal plots show
the univariate posterior parameter distributions, off-diagonal plots show the bivariate scatter
plot of posterior draws. The blue line indicates the maximum a posteriori estimate of the key
recovery rate parameters from the Python MCMC based on a gamma distribution ($\alpha = 2.5$, A
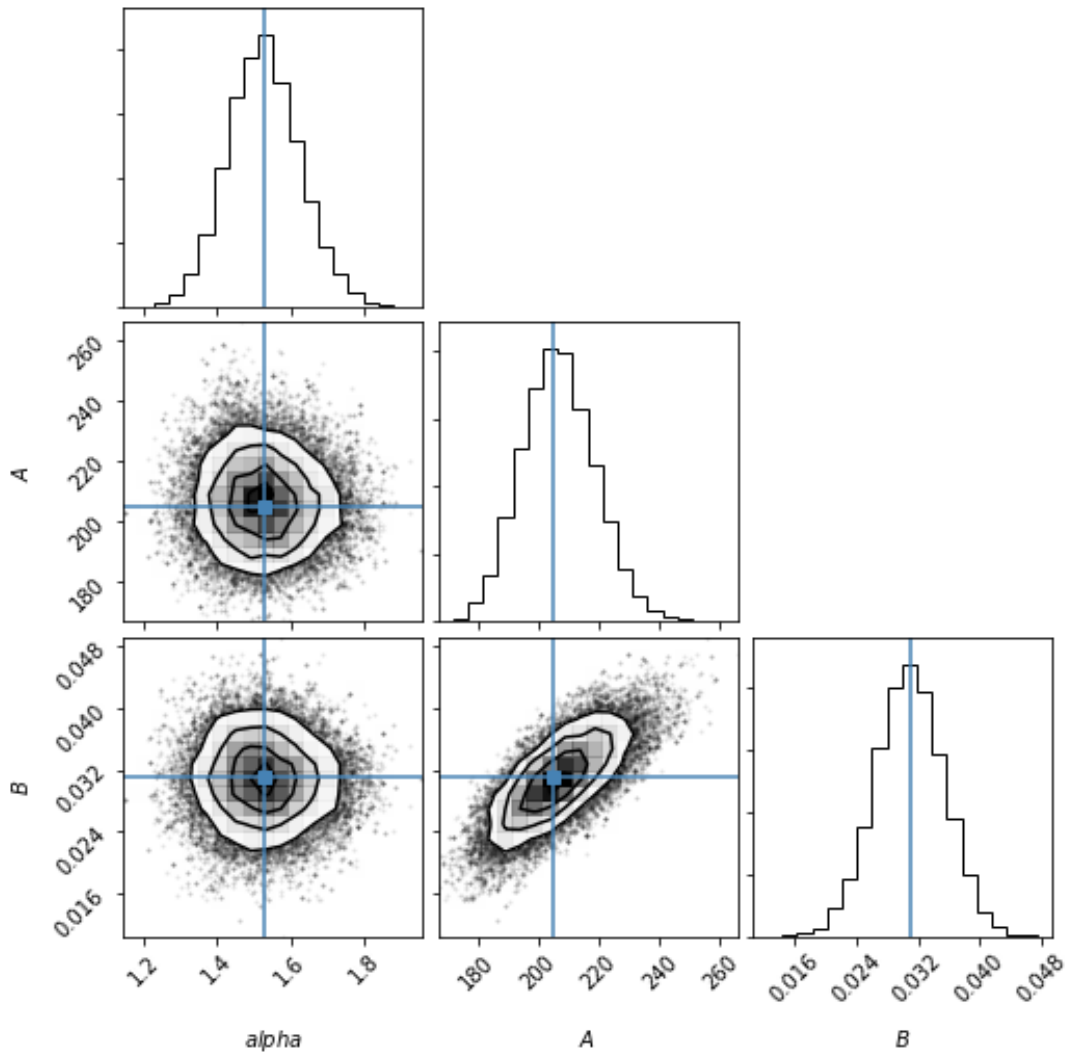= 55, B = 0.002), resulting in a maximum carriage duration of 87 days.

Figure 5.6: Scatter plot matrix of parameter distributions for serotype 19F. Diagonal plots
show the univariate posterior parameter distributions, off-diagonal plots show the bivariate
scatter plot of posterior draws. The blue line indicates the maximum a posteriori estimate of
the key recovery rate parameters from the Python MCMC based on a gamma distribution (α
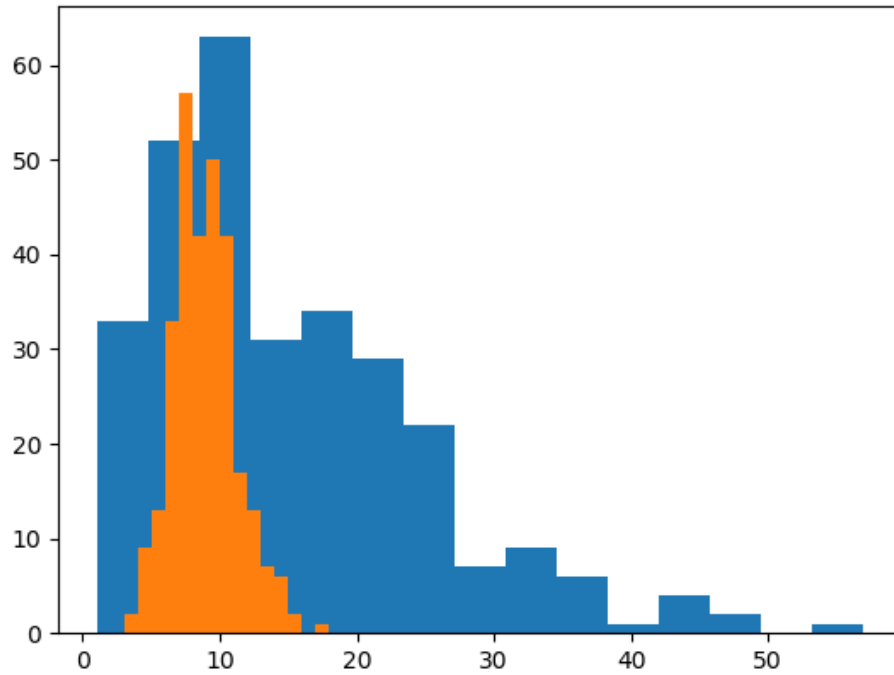= 1.53, A = 205, B = 0.031), resulting in a maximum carriage duration of 214 days.

Figure 5.7: Distribution of total previous colonisations in data (blue) versus in the simulation with the naïve population (orange) over two years.

5.3.2 Population stability

To check the impact of population size on model predictions, pre-vaccination simulations with 500, 1000, 5000 and 10,000 people in the French population were run and resulting carriage prevalence compared across age groups. The full 10,000-person model took around four weeks to run completely, whereas the 5000-person model took about two weeks; the 1000-person model took five days, and the 500-person model took less than two days.

These were initially run with default parameters ($\varepsilon = 0.1, \theta = 0.1, \sigma = 0.4$), where carriage prevalence was found to be very low (overall carriage prevalence 10%). This may be because immunity from previous colonisation was not protective enough to reduce carriage episode length or to prevent further carriage acquisition. Regardless, based on the age-specific carriage prevalence, the populations appear relatively stable over several simulations, although full

Table 5.5: Pre-vaccination carriage prevalence across different population sizes (N) and age groups for a French setting over n number of simulations with the default parameter set ($\varepsilon = 0.1, \theta = 0.1, \sigma = 0.4$). Carriage prevalence values for the 5000-individual population indicate the mean with 95% confidence interval over n simulations. Running time indicates the running time for one simulation.

N	n	Pre-PCV carriage prevalence				Running time
		Overall	Child	Adult	Elder	
500	1	12.2%	9.4%	14.6%	4.8%	33 hrs
1000	1	9.8%	8.3%	11.9%	2.9%	44 hrs
5000	20	13.2% (9.6% - 16.9%)	4.8% (3.7% - 6.0%)	15.4% (11.0% - 19.8%)	8.4% (6.6% - 10.3%)	290 hrs
10000	1	13.7%	4.0%	15.6%	9.7%	431 hrs

convergence to a stable set of values does require larger populations (Table 5.5). The results do suggest, however, that the model can be run using the less resource- and time-intensive 5000-person population with a fair degree of accuracy. As such, the rest of the analyses reported in this chapter are based solely on the 5000-individual simulations.

5.3.3 Vaccine design

Initial vaccine designs (hPCV13) were considerably different across the vaccine strategies, with only serotypes 5, 7F, 14 and 19A in common between the three vaccine strategies (Table 5.6). Unsurprisingly, the disease incidence and local invasiveness strategies had similar hPCV13, as the serotypes causing the most disease would be more likely to be those that are most invasive. Notably, serotype 4 was missing in both country-specific strategies whereas it was highly invasive in the global dataset and included in the first formulation in the global strategy. The global invasiveness strategy initially included five of the seven PCV7 serotypes, all three additional PCV10 serotypes, and one additional PCV13 serotype. Serotypes 8 and 22F, also included in hPCV13, will be included in PCV20.

Overall, however, vaccine strategies yielded inclusion of mostly similar serotypes by hPCV30 (Table 5.6). Serotypes 1, 3, 5, 6B, 7F, 9N, 10A, 12F, 14, 16F, 17F, 18C, 19A, 19F, 23F, 24F, 31, and 33F were included in both invasiveness strategies and in a majority of the disease incidence

strategy simulations. Nine of these serotypes are currently included in the current licensed PCV formulations. Serotype 3, which was highly invasive in the local invasiveness datasets of both countries, was not included in hPCVs until hPCV30 in the global invasiveness strategy. This serotype has been a concerning cause of disease in reality as PCV13 is not as effective against it as against the other serotypes.

There did not appear to be any association between serotype inclusion and the duration of carriage, as the targeted serotypes represent a variety of carriage durations that range from the absolute minimum (serotype 31, 6 days) to the absolute maximum (serotype 15B/C, 482 days) (Table 5.4, Figure 5.8). This suggests that the ability to cause disease locally or globally is not strongly related to carriage duration or fitness in colonisation. In the case of the disease incidence strategy, it may also indicate that serotype fitness may also be a result of differing effective contact rates, which here were kept constant across serotypes.
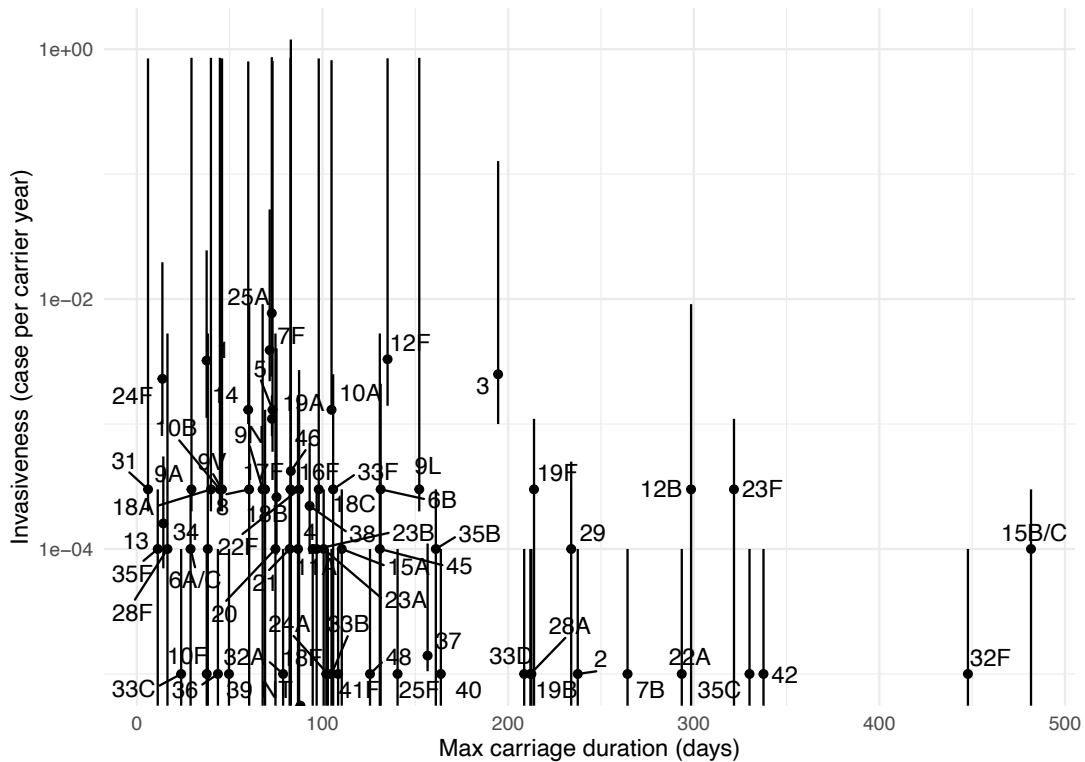

Figure 5.8: Association between maximum carriage duration and invasiveness values from French dataset.

The serotypes most frequently included in hPCVs in France in the disease incidence strategy

Table 5.6: Serotypes included in each vaccine strategy. In the case of the disease incidence
strategy, the most frequently included serotypes over the default scenario simulations were
included in the table. VT7: serotypes included in PCV7. VT10: serotypes included in PCV10.
VT13: serotypes included in PCV13. NVT: serotypes not included in currently licensed
vaccines.

Vaccine strategy	hPCV13	hPCV20 (additional)	hPCV30 (additional)	hPCV30 (all)
Disease incidence	5, 7F, 9A, 10A, 10B, 12F, 14, 16F, 18C, 19A, 19F, 24F, 25A	1, 3, 6B, 17F, 18A, 22F, 23F, 33F	8, 9L, 9N, 15B/C, 24A, 28F, 29, 31, 35B, 35F, 46	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 10A, 10B, 12F, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23F, 24A, 24F, 25A, 28F, 29, 31, 33F, 35B, 35F, 46
Local invasiveness	1, 3, 5, 7F, 10A, 10B, 12F, 14, 16F, 19A, 24F, 25A, 46	9A, 9L, 9V, 12B, 18A, 18C, 31	6B, 8, 9N, 17F, 18B, 19F, 22F, 23F, 33F, 38	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 14, 16F, 17F, 18A, 18B, 18C, 19A, 19F, 22F, 23F, 24F, 25A, 31, 33F, 38, 46
Global invasiveness	1, 4, 5, 6B, 7F, 8, 9V, 14, 18C, 19A, 22F, 27, 38	6A, 9N, 12F, 13, 17F, 18B, 19F, 20, 23F, 33F	3, 10A, 15A, 15B/C, 16F, 23A, 23B, 24F, 31, 35B	1, 3, 4, 5, 6A, 6B, 7F, 8, 9N, 9V, 10A, 12F, 13, 14, 15A, 15B/C, 16F, 17F, 18B, 18C, 19A, 19F, 20, 22F, 23A, 23B, 23F, 24F, 27, 31, 33F, 35B, 38

indicated that the differing scenarios did not have a large impact on which serotypes were
targeted (Table 5.7). There was no consistent trend between serotypes that were or were
not included frequently between scenarios, as serotype-specific models may be susceptible
to a high degree of noise. For example, even though altering the non-specific immunity
(ε) alters the recovery rate, one might expect that with increasing non-specific immunity,
the less fit serotypes would be more frequently included in the vaccines. However, when
looking at the association between maximum carriage duration and French invasiveness,
many of the serotypes chosen frequently in the disease incidence strategy across scenarios
(such as serotypes 3, 12F and 25A) had the highest invasiveness rather than a slower recovery
rate. Interestingly, serotype 6A/C was only in the top thirty most frequent serotypes in
the 85% vaccine efficacy scenario. Serotype 6A is included in PCV13 as it was responsible
for a substantial amount of disease post-PCV7, particularly in Asian and African countries
[199, 200, 440].

5.3.4   Carriage prevalence across scenarios

Vaccines were varied in how successful they were against their respective targets post-
vaccination in the default scenarios (Figure 5.9). The average carriage prevalence of VTs was
extremely low across age groups and vaccination periods, rarely going above 15% in the case
of the disease incidence strategy and less than 2% in most cases in the local invasiveness and
global invasiveness strategy simulations. The wide confidence interval bands across most
age groups under the disease incidence strategy may be due to the inclusion of different
serotypes in each simulation which will exhibit different carriage duration properties and
therefore have varying effects on the carriage dynamics of the population.

Under the local and global invasiveness strategies, average VT carriage prevalence was
extremely low across age groups. Children showed a non-significant increase in average
carriage prevalence across VTs and vaccine periods. Despite this, there was a decrease in
average VT13 carriage prevalence with vaccination in adults and the elderly, indicating

that the vaccines worked and potentially suggesting herd immunity. One potential reason
for this discrepancy between children and adults is that the time between vaccines is 10
years, whereas children spend at most 5 years in this age group and must wait at least 2
to 3 years before they are vaccinated in the model, so the benefits of vaccination are likely
to be seen once they age out of their age category. VT20 also decreased across vaccination
periods even before they were being targeted by the vaccines. This was the case with VT30
as well, although this serotype group strangely increased after hPCV30, indicating that the
hPCV30 was not successful in removing these serotypes from circulation. Upon further
inspection, these results were driven by one simulation that had extremely high carriage
prevalence of VT30 post-hPCV20 and hPCV30 (80% and 60% respectively). Other simulations
showed successful removal of these serotypes as VT30 carriage prevalence was below 1%
post-hPCV30 in all other simulations. Interestingly, the additional serotypes (VT20 and
VT30) did not increase in carriage prevalence after successive vaccinations targeting VT13,
suggesting that there were other highly prevalent serotypes in carriage and minimal serotype
replacement taking place under this vaccination strategy.

Table 5.7: Most frequently included vaccine serotypes for each disease incidence strategy sensitivity scenario in France with 5000 individuals in the population. VT7: serotypes included in PCV7. VT10: serotypes included in PCV10. VT13: serotypes included in PCV13. NVT: serotypes not included in currently licensed vaccines. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

| Model description	No. simulations	Most frequent hPCV vaccine serotypes
$\theta = 0.2$	19	1, 3, 5, 6B, 7F, 8, 9A, 9N, 9V, 10A, 10B, 12B, 12F, 13, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23F, 24F, 25A, 31, 33F, 38, 46
$\theta = 0.5$	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 13, 14, 15B/C, 16F, 18A, 17F, 18C, 19A, 19F, 23F, 24F, 25A, 31, 33F, 35F, 46
$\sigma = 0.6$	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23F, 24F, 25A, 31, 33F, 38, 46
$\sigma = 0.8$	20	1, 3, 5, 6B, 7F, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 13, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23F, 24F, 25A, 31, 33F, 35F, 46
$\varepsilon = 0.25$	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 10A, 10B, 12B, 12F, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23B, 23F, 24F, 25A, 28F, 31, 33F, 46
$\varepsilon = 0.5$	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 14, 15B/C, 16F, 18A, 18B, 18C, 19A, 19F, 22F, 23F, 24F, 25A, 31, 33F, 35F, 46
$v = 75\%$	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 14, 15A, 16F, 18A, 18C, 19A, 19F, 22F, 23B, 23F, 24F, 25A, 33F, 35F, 38, 46
$v = 85\%$	19	1, 3, 5, 6A/C, 6B, 7F, 8, 9A, 9L, 9N, 10A, 10B, 12B, 12F, 14, 16F, 17F, 18A, 18C, 19A, 19F, 21, 22F, 23F, 24F, 25A, 31, 33F, 34, 46
$p = 5$ years	20	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9N, 9V, 10A, 10B, 12B, 12F, 14, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 22F, 23B, 23F, 24F, 25A, 31, 33F, 46
$p = 25$ years	19	1, 3, 5, 6B, 7F, 8, 9A, 9L, 9V, 10A, 10B, 12B, 12F, 14, 15A, 15B/C, 16F, 17F, 18A, 18C, 19A, 19F, 23B, 23F, 24F, 25A, 31, 33F, 35F, 46
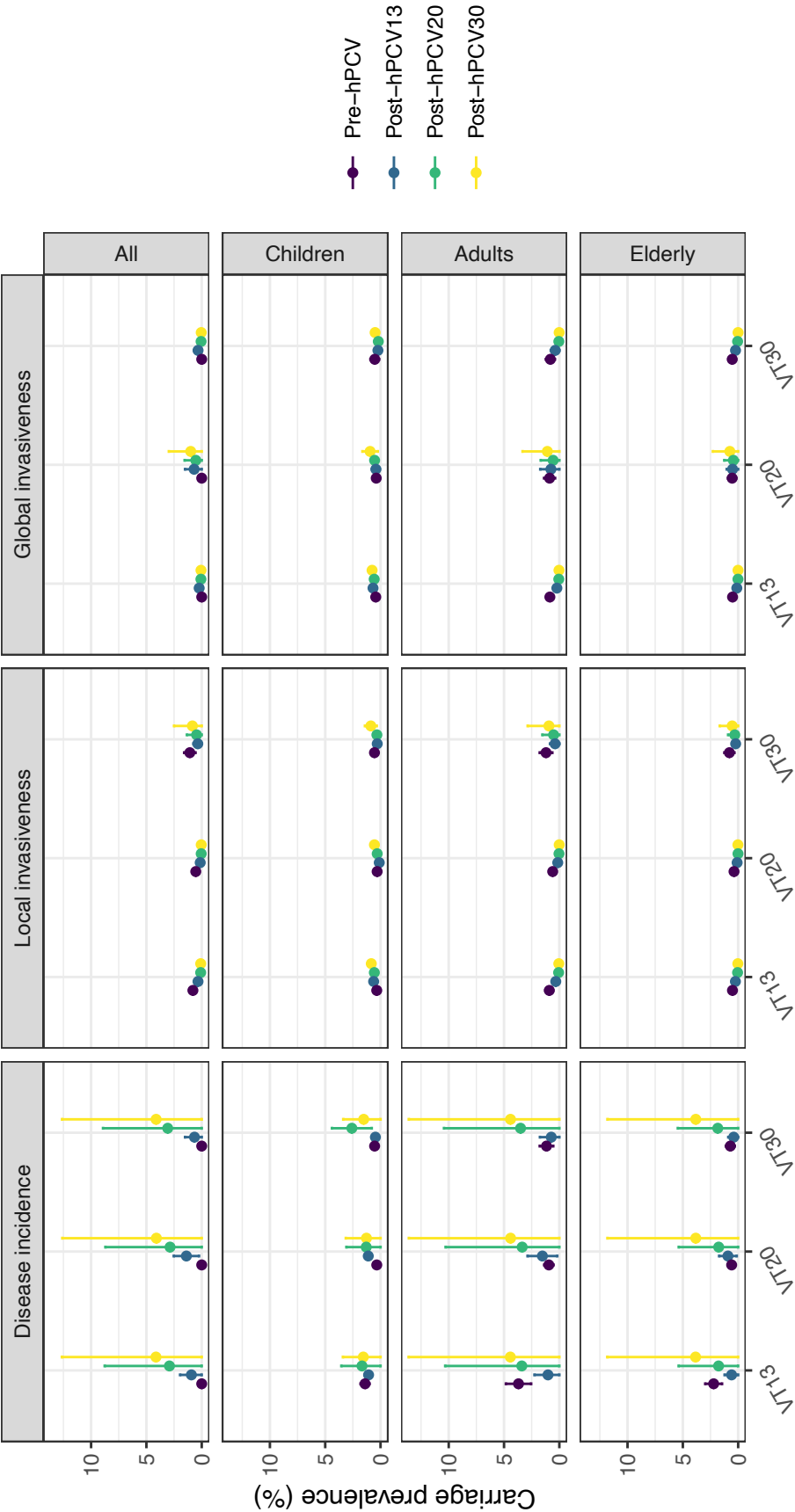
Figure 5.9: Average carriage prevalence of VTs in the 5000-person populations in the French setting with the default parameter set. VT13 indicates the serotypes included in hPCV13, VT20 indicates the additional serotypes included in hPCV20, and VT30 indicates the additional serotypes included in hPCV30.

Overall pre-vaccination carriage prevalence (including all serotypes) was not altered significantly across most scenarios in France (Figure 5.10). Carriage prevalence was found to be quite low across age groups and scenarios, being less than 20% in children and elderly in all scenarios and less than 30% in most scenarios in adults. This may be due to the closed population, as individuals cannot die from their carriage (or disease) episodes in the model and thereafter be replaced by new susceptible individuals. This allows individuals to build natural immunity (both specific and non-specific) that does not wane in this model, and therefore may result in a lower carriage prevalence [254]. This is shown in the specific immunity scenarios where σ was set higher (0.6, 0.8) than the default value (0.4), and results in significantly lower average carriage prevalence values ($<$10%) across age groups compared to all other. Children consistently had a significantly lower carriage prevalence among all age groups. This may be due to children under five years making up the smallest age group (6% of the population in both France and the US), and therefore not being as likely to be randomly selected in the stochastic model with the chosen parameters despite their higher rates of infection. Though if this were the case, one would expect the carriage prevalence to be proportional to their smaller population. More likely this is due to the contact matrices, as children under 5 years have much fewer contacts with the general population than the adult and elderly group. The vaccine efficacy and duration scenarios should not impact the pre-PCV carriage prevalence, as demonstrated by the results with non-significant differences between these and the default scenario across age groups. Unsurprisingly, the average carriage prevalence was not significantly different between France and the US, as these settings differ only slightly in contact matrices but not in the epidemiological parameters used.
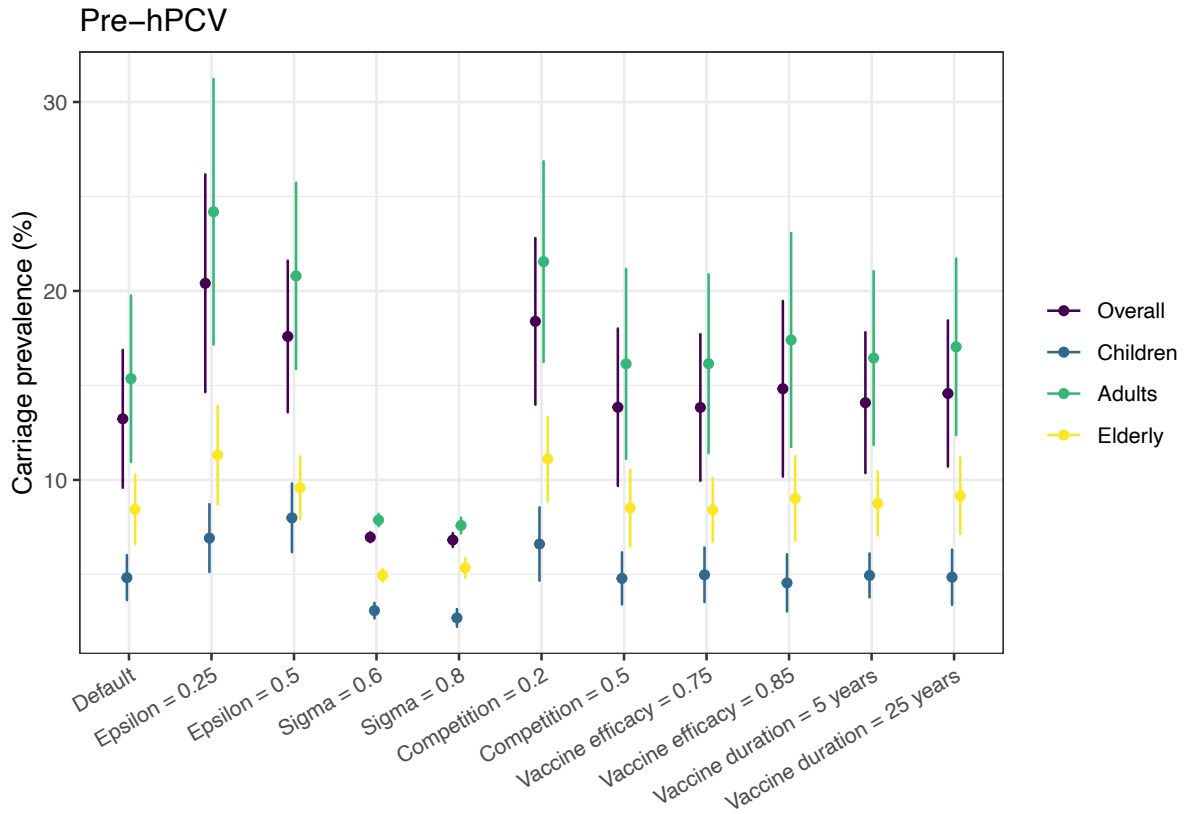
Figure 5.10: Average pre-vaccination carriage prevalence across scenarios (varying each parameter respectively) run twenty times each with the 5000-person populations in French setting. Default parameter set: $\varepsilon = 0.1, \theta = 0.1, \sigma = 0.4$. ε: non-specific immunity, σ: specific immunity, θ: competition.

Interestingly, average carriage prevalence increased in all age groups post-vaccination in the default scenario, almost doubling between post-hPCV13 and post-hPCV20 (Table 5.8). This increase in carriage prevalence over vaccination periods may indicate serotype replacement under the default parameter set, as VT are vaccinated against and leave an open niche for new serotypes against which community-wide specific immunity has not been built or which may be too weak. However, since average VT prevalence was predicted to be low across the different vaccination strategies (Figure 5.9), this may indicate that specific immunity is too weak.

Post-vaccination carriage prevalence, like in pre-vaccination carriage prevalence, was lowest in children compared to other age groups and was not significantly different between vaccine

Table 5.8: Post-vaccination carriage prevalence across different strategies and age groups for the French setting with the default parameter set (ε=0.1, $\theta = 0.1$, σ=0.4, $v = 95\%$, $\omega = 10$ years). Results indicate the mean with 95% confidence interval over n simulations of the 5000-individual population.

Strategy	Overall	Children	Adults	Elderly
Post-PCV13				
Disease incidence	35% (25.9% - 44.2%)	10.9% (8.5% - 13.2%)	40.7% (30.2% - 51.2%)	23.3% (16.7% - 29.9%)
Local invasiveness	36.7% (28.1% - 45.3%)	10.9% (8.5% - 13.3%)	42.7% (32.9% - 52.6%)	24.2% (17.8% - 30.5%)
Global invasiveness	34.1% (24.9% - 43.3%)	10% (7.6% - 12.5%)	39.6% (29.1% - 50.2%)	22.9% (16.2% - 29.6%)
Post-PCV20				
Disease incidence	61% (52.3% - 69.7%)	17.5% (14.4% - 20.5%)	67.9% (58.9% - 76.9%)	50.9% (41.3% - 60.6%)
Local invasiveness	62% (53.1% - 70.8%)	17.5% (14.8% - 20.3%)	68.8% (59.6% - 78%)	52.2% (42.6% - 61.8%)
Global invasiveness	58.1% (48% - 68.2%)	16.7% (14% - 19.3%)	64.7% (54% - 75.5%)	48.4% (37.9% - 58.9%)
Post-PCV30				
Disease incidence	78.9% (73.7% - 84.1%)	24.5% (21.5% - 27.5%)	84.8% (80% - 89.5%)	75.3% (67.9% - 82.7%)
Local invasiveness	77.5% (69.9% - 85.1%)	24.4% (21.3% - 27.5%)	83% (75.3% - 90.7%)	74.9% (65.8% - 84%)
Global invasiveness	75.4% (67.5% - 83.3%)	23.8% (20.8% - 26.9%)	81.2% (73.3% - 89%)	71.1% (61.2% - 81%)

strategies (Figure 5.11, Figure 5.12, Figure 5.13). Increasing the non-specific immunity resulted in an increased carriage prevalence in children, though non-significant in all cases. Increasing the non-specific immunity reduces fitness differences and enhances diversity [57]. Consistent with the pre-vaccination results, the specific immunity scenarios where σ was elevated resulted in extremely low average carriage prevalence (<10%) across age groups and stayed consistently low with each round of vaccinations. Changing the competition parameter had a non-significant effect on average carriage prevalence across age groups, although there was a noticeable decrease in carriage prevalence with increased competition across age groups in France post-hPCV13 (Figure 5.11). One reason for this may be that a larger competition parameter dampens the serotype-specific FOI, and therefore enhances serotype diversity. This may promote a larger overall carriage prevalence up to a certain extent (as seen with $\theta =$ 0.2) until competition may become too high ($\theta = 0.5$), and eventually shifting the transmission dynamics to a mono-carriage system. How levels of competition impact carriage prevalence and serotype diversity merits further exploration. Decreasing the vaccine efficacy also did not have a significant effect on the carriage prevalence in any age group or under any strategy regardless of hPCV period (Figure 5.11, Figure 5.12, Figure 5.13), which is consistent with real-world findings that carriage prevalence does not change post-vaccination as new serotypes fill the empty niche left by the vaccinated serotypes once they are removed from circulation [441]. Surprisingly, the vaccine duration of protection did not alter the carriage prevalence across vaccine periods, even at its lowest (Figure 5.11, Figure 5.12, Figure 5.13). This may be due to the strength of the natural immunity parameters and suggests that further parameter exploration around the immunity parameters is desirable. However, assessing the strength of immunity is very difficult to measure in exposed populations in the absence of experimental studies

In some simulations of both global and local invasiveness strategies, a single low-invasiveness serotype was able to gain a competitive advantage and then was never vaccinated against as only the highly invasive types were targeted in this strategy (Figure 5.14). In this case, specific immunity may be too weak to overcome multiple colonisations of the same type, as
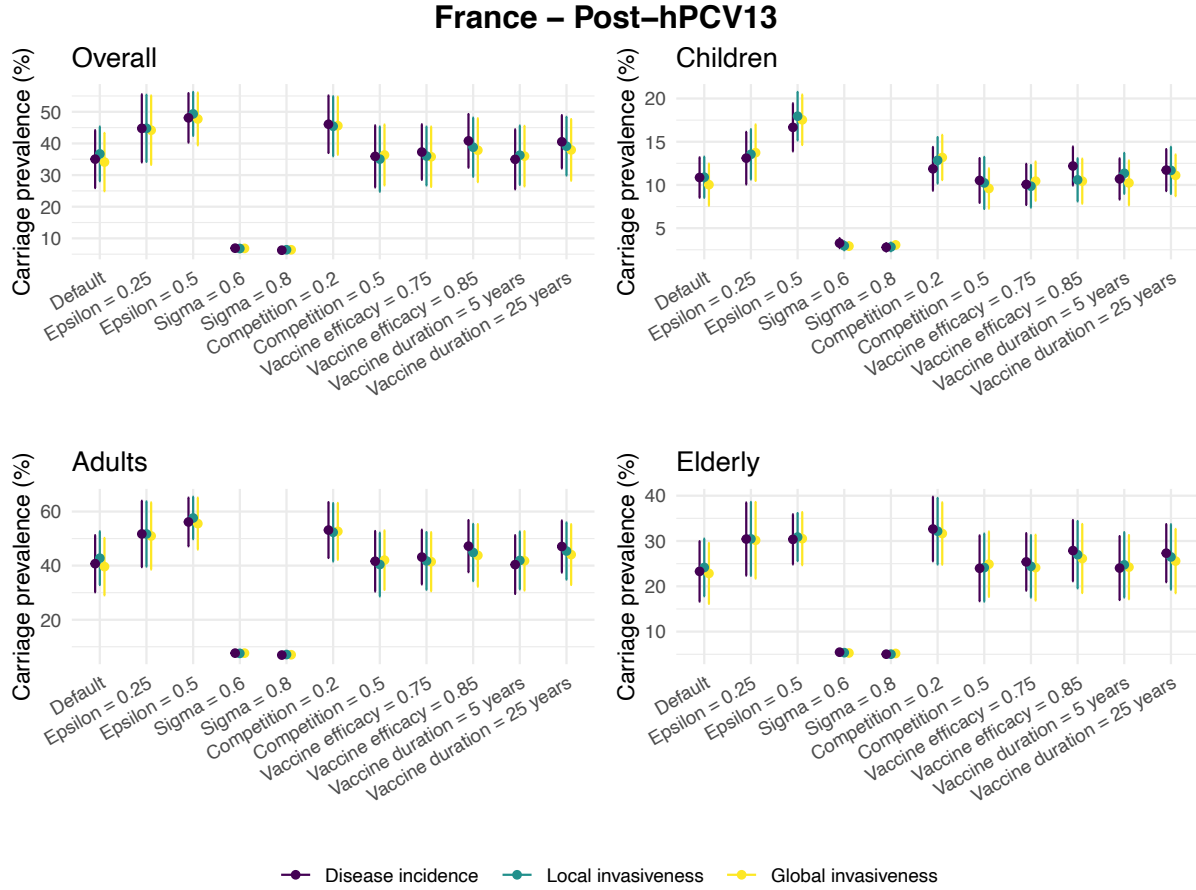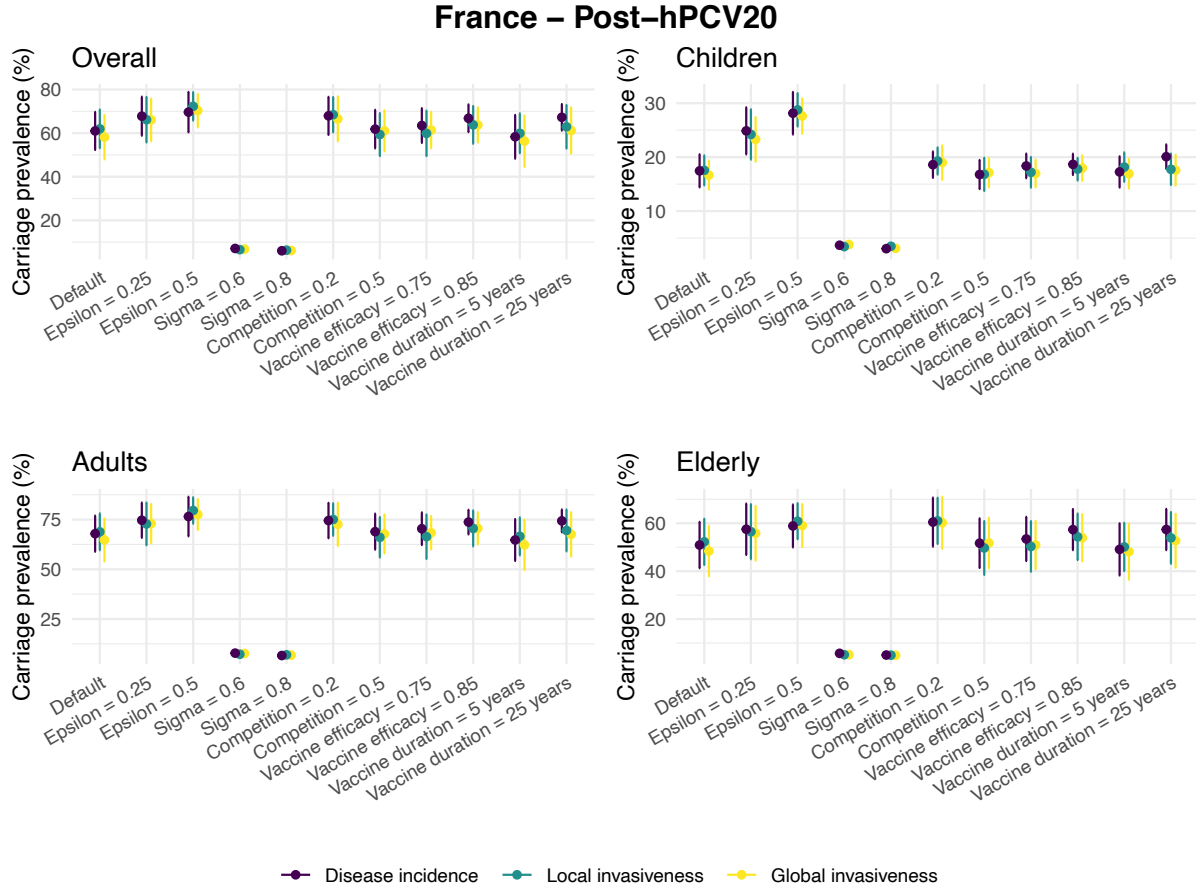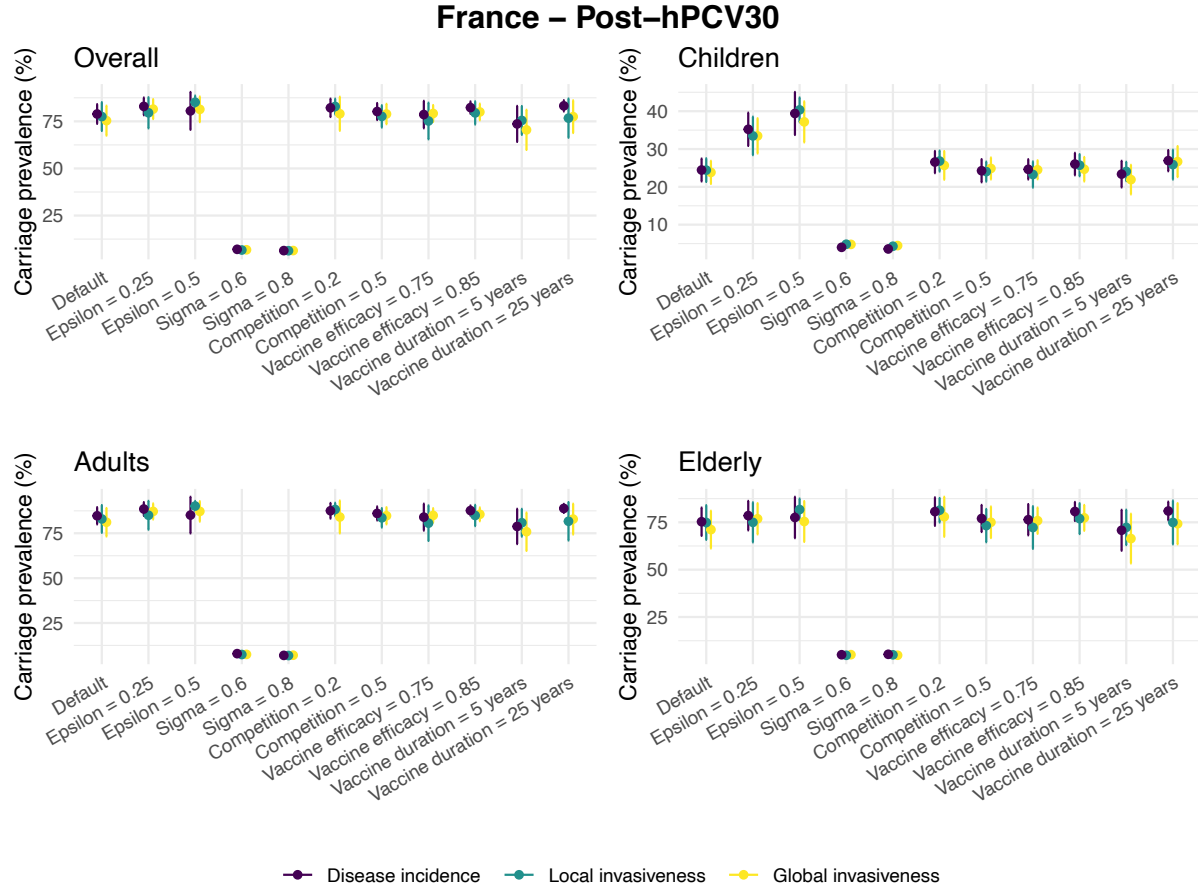
Figure 5.11: Average post-hPCV13 carriage prevalence across scenarios (varying each parameter respectively) run with the 5000-person population in France. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%, \omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.12: Average post-hPCV20 carriage prevalence across scenarios (varying each parameter respectively) run with the 5000-person population in France. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%, \omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.13: Average post-hPCV30 carriage prevalence across scenarios (varying each parameter respectively) run with the 5000-person population in France. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%, \omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

it is independent of the number of specific previous colonisations. Individuals perpetually colonised by this type would contribute to the serotype's higher and increasing FOI, allowing the serotype to continue proliferating. While the serotype is unlikely to cause as much disease as one that is highly invasive, this situation may become problematic in reality if the serotype in question has a high case-fatality rate.
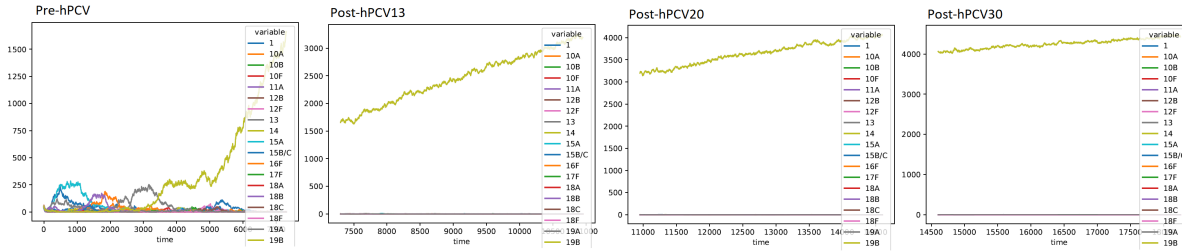
Figure 5.14: Illustrative example of local invasiveness strategy simulation with 5000 individuals in French setting where a non-highly invasive serotype (in this case nontypeable pneumococci) proliferates undeterred by vaccination with other serotypes present at extremely low prevalence. Legend cut off for legibility purposes.

5.3.5 Incidence rate ratios after expanding serotype coverage in vaccination

Pooled incidence rate ratio (IRR) estimates fitted from a fixed effects model to estimate the effect size after vaccination over the simulations demonstrated comparable, mostly non-significant results in the decreased incidence before and after vaccination with specific hPCVs across strategies and age groups (Figure 5.15). Although hPCVs targeted children between the ages of two and three, they showed a reduced, significant effect in the adult and aggregate age groups post-hPCV13 under the disease incidence strategy. Narrower confidence intervals in these two age groups are likely due to the larger size of these age groups compared to children and the elderly. Pooled estimates across age groups were therefore driven by adults.

While pooled IRRs showed non-significant effects across most age groups, vaccines and strategies, average incidence rates were significantly reduced post-vaccination in the adult and aggregate age groups between pre-vaccination and post (Figure 5.16), and were reduced

(though non-significantly) in the elderly as well. The average disease incidence in children appeared to increase post-vaccination. There were no significant reductions in disease after vaccination with hPCV13, and this was consistent regardless of strategy. All strategies seemed equally successful at reducing the average disease incidence. The discrepancy between the pooled IRR results and the average incidence are likely due to higher sensitivity of the fixed effects model, as this takes the weighted average and considers IRR confidence intervals, compared to a crude averaging of incidence rates. Regardless, these results suggest that the vaccines appeared to be effective against disease, and that running the model with a larger population or bootstrapping would likely result in narrower confidence intervals in the average incidence rates and pooled IRRs, as the sample would be larger.
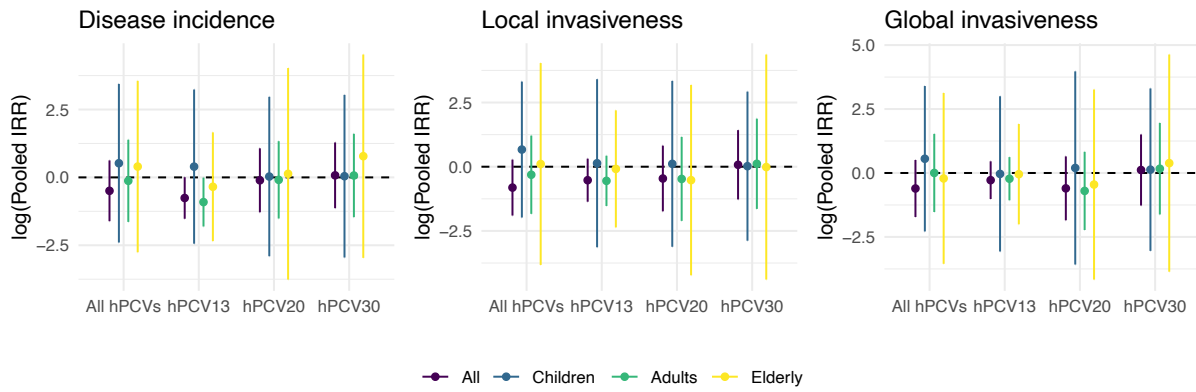
Figure 5.15: Comparison of the natural log of the pooled incidence rate ratios across vaccination strategies and age groups in French setting under the default scenario with a 5000-individual population. Dashed line indicates where the pooled incidence rate ratios are 1. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

There was no significant difference when comparing the IRRs across sensitivity scenarios within each age group under the local invasiveness (Figure 5.17), global invasiveness (Figure 5.18), or disease incidence vaccine strategies (Figure 5.19). Strategy results were mixed, with most showing non-significant effect sizes, and some scenarios showing non-significant reductions in IPD incidence in the aggregate age group and the adult age group. In these two age groups, hPCV13 was the only vaccine that significantly reduced disease in some scenarios and hPCV30 was least successful in reducing disease across strategies and scenarios.
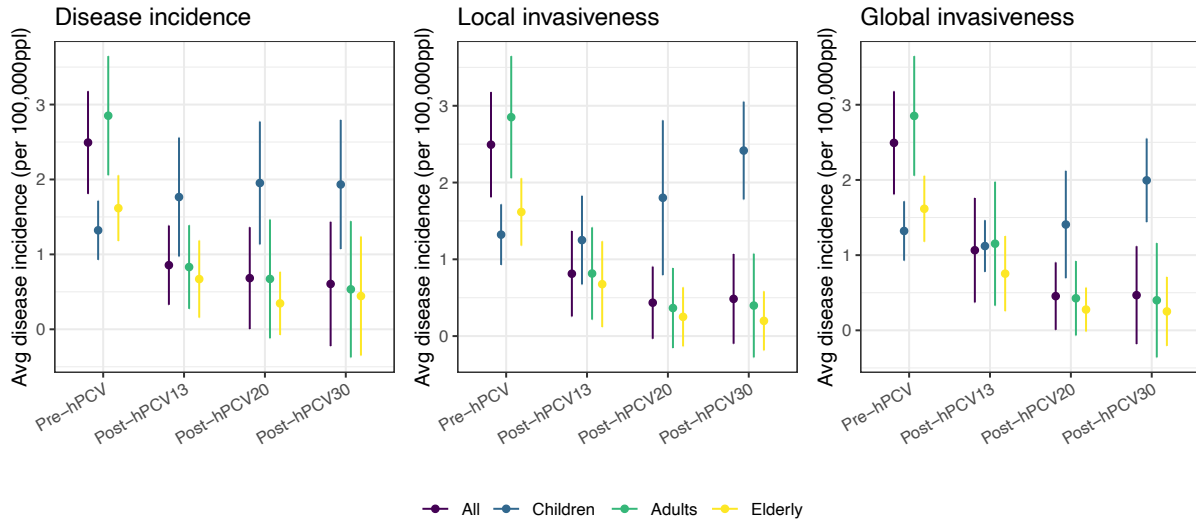
Figure 5.16: Comparison of average disease incidence across vaccination strategies and age groups in French setting under the default scenario with a 5000-individual population. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

This is likely due in part to the serotypes targeted also having the highest invasiveness estimates (Table 5.6, Table 5.7) such that removing these in the first hPCV formulation will result in less invasive potential in the remainder of the pneumococcal population. Pooled IRRs after hPCV13 were significantly less than 0 in both age groups and across strategies when vaccine efficacy was 85%. There were notable differences between the strategies with the same scenario, e.g., a lower vaccine efficacy of 75% resulted in significantly reduced IRRs in both age groups under the disease incidence strategy post-hPCV13 (Figure 5.19) but was not the case in the other strategies (Figure 5.17, Figure 5.18). In line with the average disease incidence rates observed in the default scenario (Figure 5.16), vaccination did not reduce IRRs in children across scenarios and strategies, emphasising the need for age-specific transmission parameters to better reflect reality. Comparing pre-vaccination and post-hPCV30, there were significant reductions in disease incidence in the aggregate age group across strategies when non-specific immunity was elevated or vaccine parameters were altered, i.e. reduction in vaccine efficacy. This is interesting as stronger non-specific immunity indicates greater serotype diversity, which one would expect to result in a larger spectrum of serotypes causing disease post-vaccination. However, a stronger non-specific immunity elicits faster recovery

rates, reducing fitness differences, and therefore prevents fitter serotypes from proliferating.
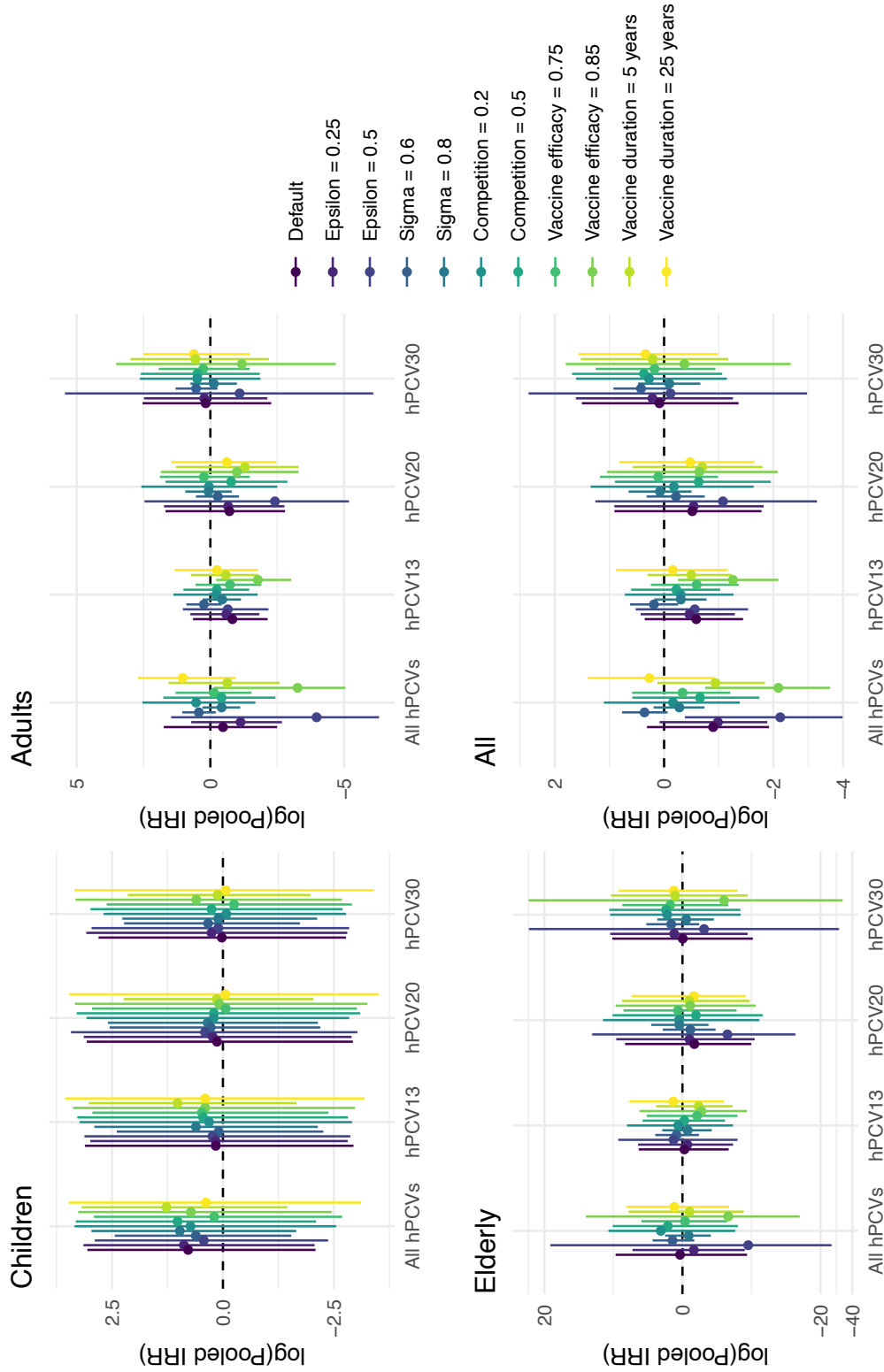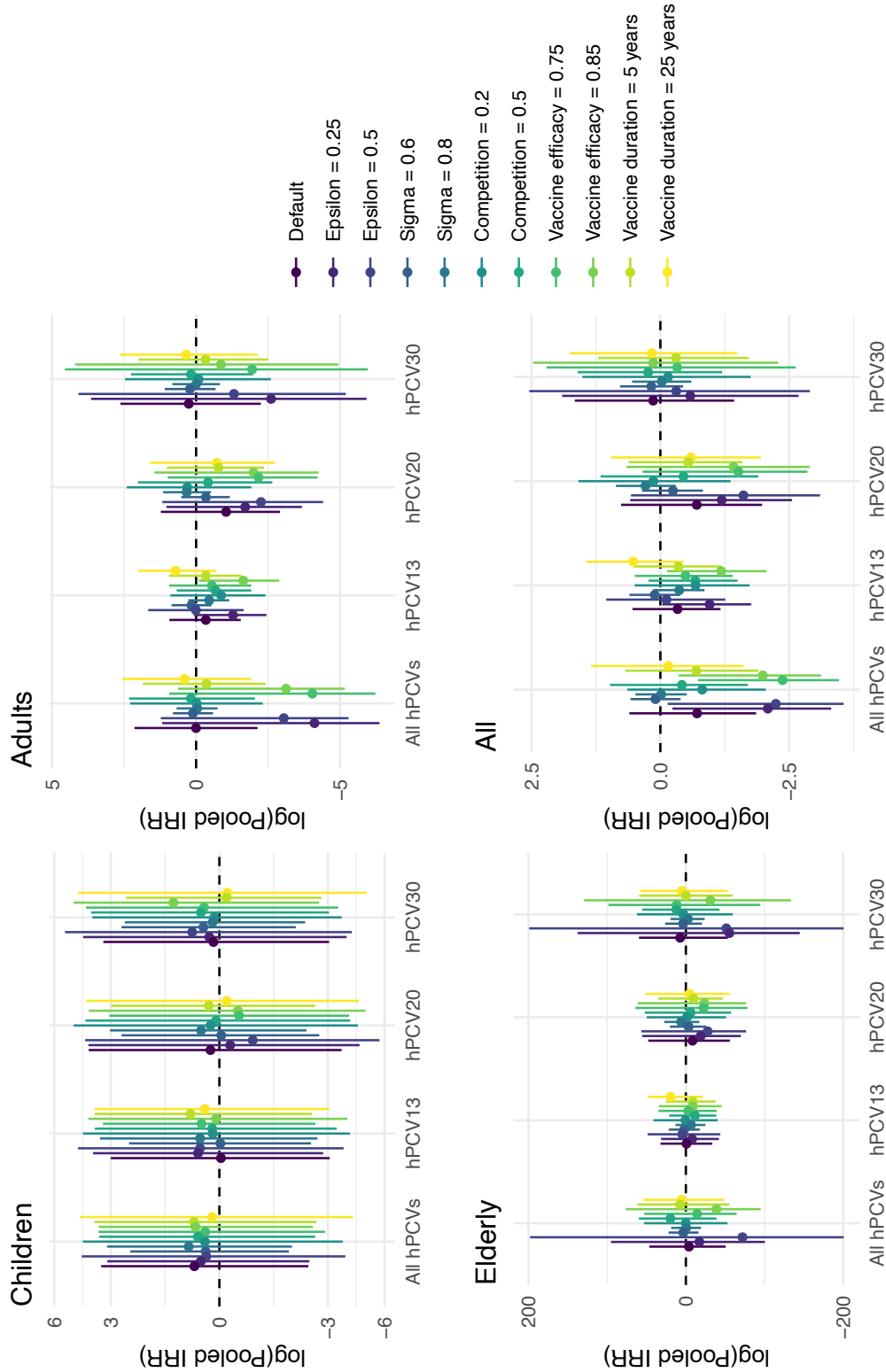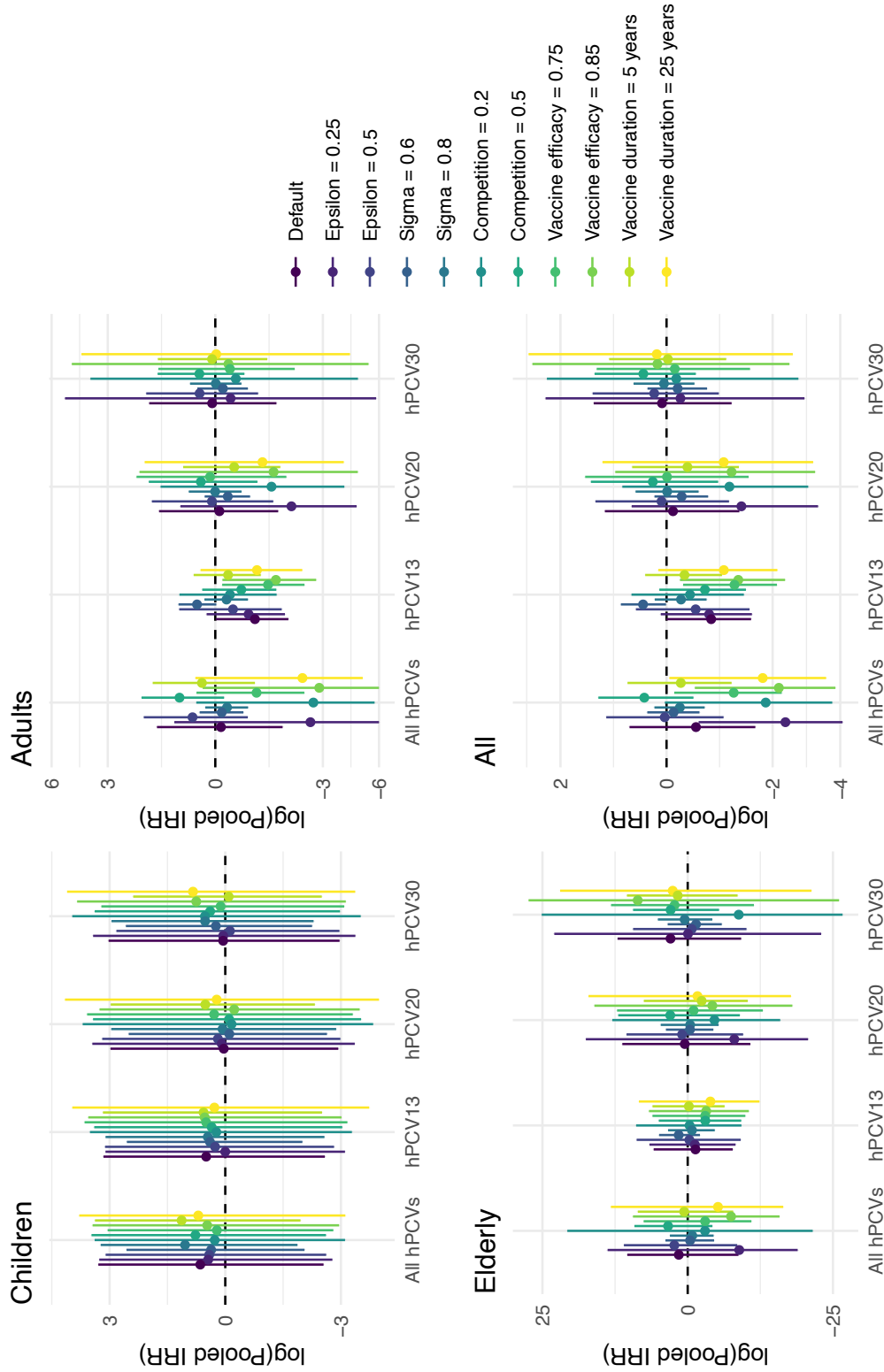
Figure 5.17: Natural log of pooled incidence rate ratios across sensitivity scenarios in each age group under the local invasiveness vaccination strategy in the French setting with the 5000-individual population. Default parameter set: $\varepsilon = 0.1$, $\sigma = 0.4$, $\theta = 0.1$, $\upsilon = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, υ: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.18: Natural log of pooled incidence rate ratios across sensitivity scenarios in each age group under the global invasiveness vaccination strategy in the French setting with the 5000-individual population. Default parameter set: $\varepsilon = 0.1$, $\sigma = 0.4$, $\theta = 0.1$, $v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.19: Natural log of pooled incidence rate ratios across sensitivity scenarios in each age group under the disease incidence vaccination strategy in the French setting with the 5000-individual population. Default parameter set: $\varepsilon = 0.1$, $\sigma = 0.4$, $\theta = 0.1$, $v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

5.3.6   Comparison across settings

The French and American settings differ by their average life expectancy, their contact matrices, and most importantly their respective invasiveness datasets. While the first two aspects are similar between the settings, the invasiveness datasets vary considerably between the two settings and show vast differences in serotypes that would be expected to have consequences in the simulations (Figure 5.20). For example, serotype 25A is highly invasive in the French setting but not in the American, whereas serotype 4 is highly invasive in the American setting but not in the French. This inclusion of serotype 4 among the highly invasive in the American dataset is similar to the global invasiveness dataset (Table 5.6). As a result of these differences between the two settings though, it is unsurprising that the average carriage prevalence pre-vaccination was non-significantly different between the two settings across scenarios (Figure 5.21), but that the two differ post-vaccination (Table 5.9). Furthermore, it provides an explanation for why the global invasiveness strategy was more successful in the US setting, as average VT carriage prevalence decreased with each vaccination across age groups except children, since the serotypes that are globally invasive were also highly invasive in the American setting (Figure 5.22) compared to the French setting.
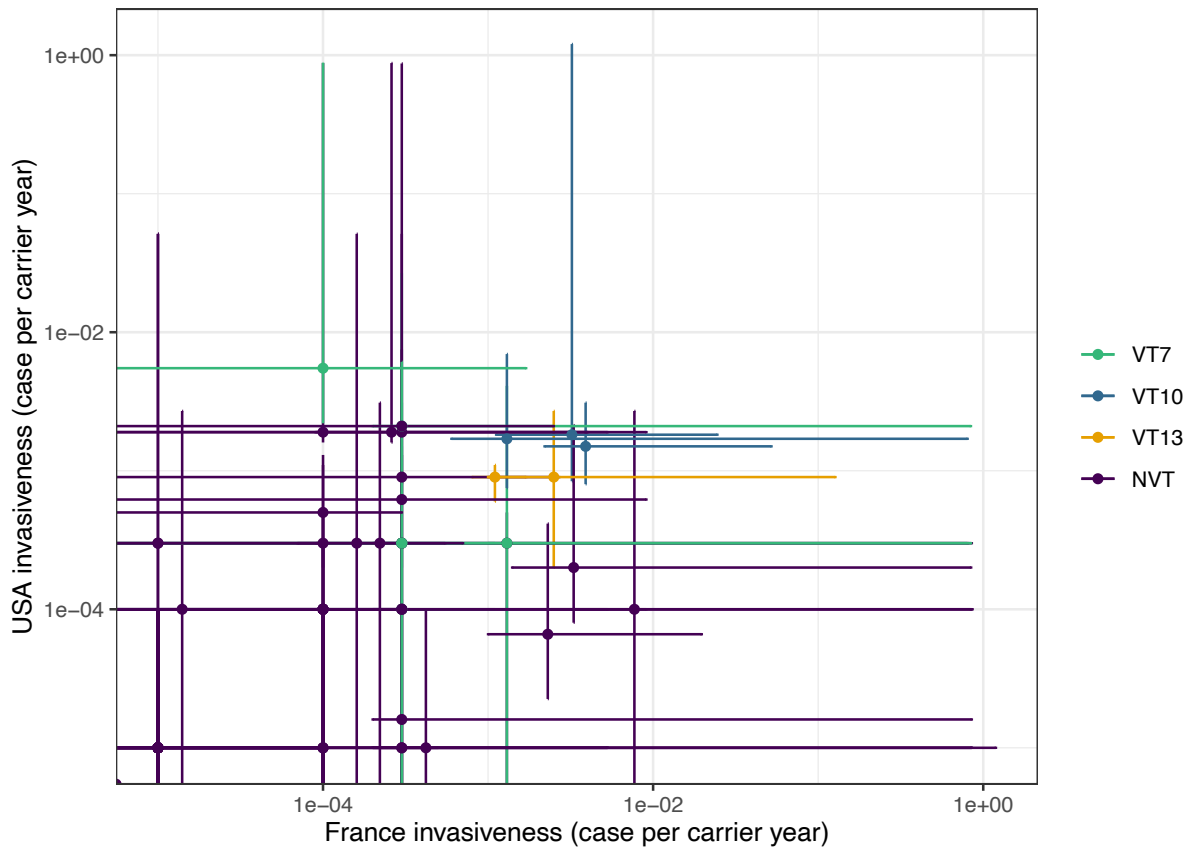
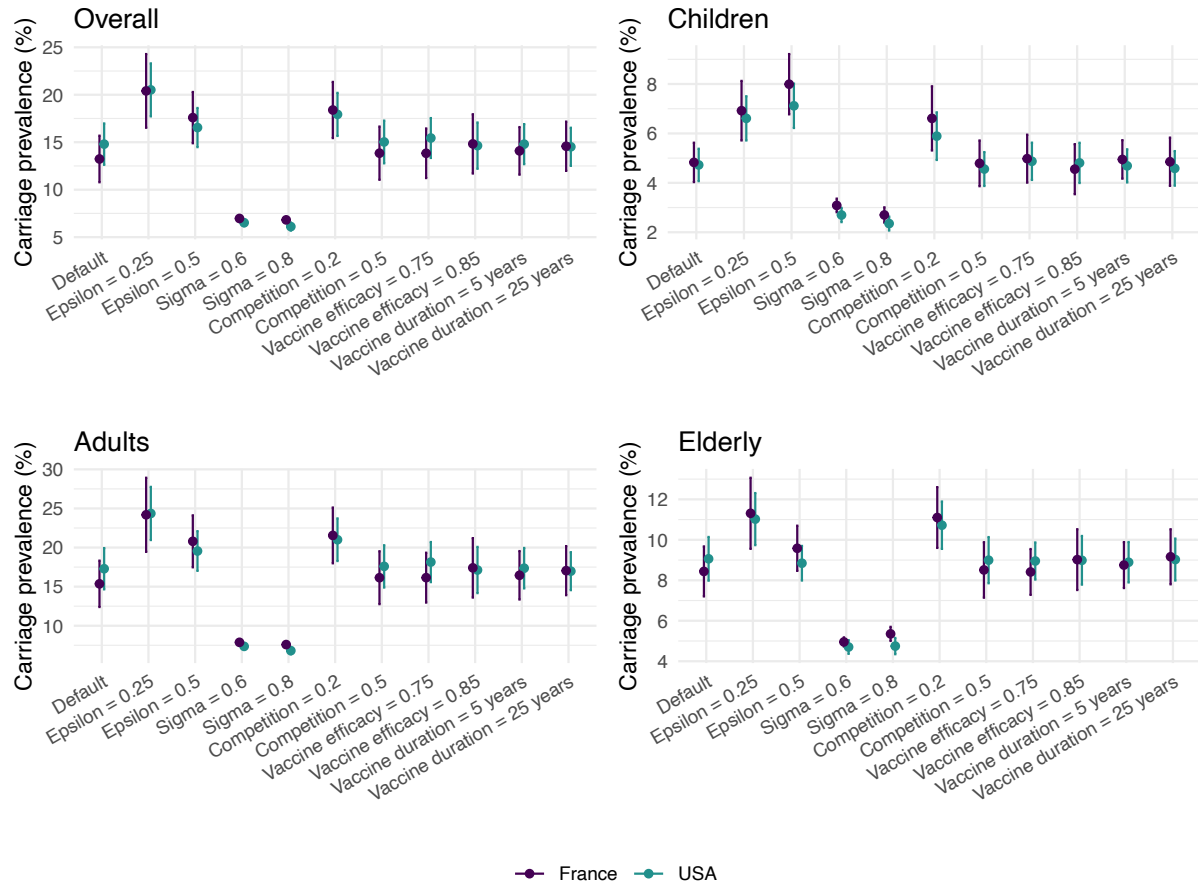Figure 5.20: Comparison of French and Massachusetts' invasiveness estimates.

Figure 5.21: Comparison of average pre-vaccination carriage prevalence across scenarios
(varying each parameter respectively) run twenty times each with the 5000-person popula-
tions in French and American settings. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v =$
$95\%, \omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine
efficacy, ω: duration of vaccine protection.

Table 5.9: Comparison of average post-hPCV carriage prevalence across age groups in French and American settings under global invasiveness strategy with the default parameter set.

Country	Overall	Children	Adults	Elderly
Post-PCV13				
France	34.1%	10%	39.6%	22.9%
	(24.9% - 43.3%)	(7.6% - 12.5%)	(29.1% - 50.2%)	(16.2% - 29.6%)
USA	36.4%	6%	41.6%	25.6%
	(29.7% - 43.1%)	(5.1% - 6.9%)	(34.2% - 49.1%)	(19.3% - 31.9%)
Post-PCV20				
France	58.1%	16.7%	64.7%	48.4%
	(48% - 68.2%)	(14% - 19.3%)	(54% - 75.5%)	(37.9% - 58.9%)
USA	62.5%	9.5%	68.3%	57.4%
	(57.3% - 67.8%)	(8.3% - 10.7%)	(63.1% - 73.5%)	(50% - 64.8%)
Post-PCV30				
France	75.4%	23.8%	81.2%	71.1%
	(67.5% - 83.3%)	(20.8% - 26.9%)	(73.3% - 89%)	(61.2% - 81%)
USA	79.2%	15.9%	84.2%	80.8%
	(76.6% - 81.8%)	(14.6% - 17.1%)	(81.5% - 86.8%)	(77.2% - 84.4%)

Figure 5.22: Average carriage prevalence of VTs across age groups in the American setting under global invasiveness strategy with default parameter set. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%, \omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Post-hPCV, American IRRs were comparable to and often non-significantly different from French ones under the global invasiveness strategy (Figure 5.23, Figure 5.24, Figure 5.25). The exception was post-hPCV13 in the 25-year vaccine protection scenario, where the American setting had a greater reduction in IRR (Figure 5.23). Disease reduction was more noticeable in the adult and aggregate age groups in the American setting than the French post-hPCV13. Other than the US setting having similar highly invasive serotypes to the global dataset, another reason for this may be that the non-specific immunity, which impacts the recovery rates and reduces fitness differences between serotypes, may allow serotypes that are highly invasive but with an average carriage duration and not included in the vaccine formulations, such as serotype 25A, to proliferate. Furthermore, pooled IRRs post-hPCV13 were significantly less than zero in the American setting in all scenarios except with the lowest

vaccine duration of protection (5 years) and vaccine efficacy (75%). After hPCV13, the point estimates of the American setting were greater than those of the French setting, though non-significant. Overall, these results highlight that further work on country comparison based on invasiveness strategies is needed to draw robust conclusions.
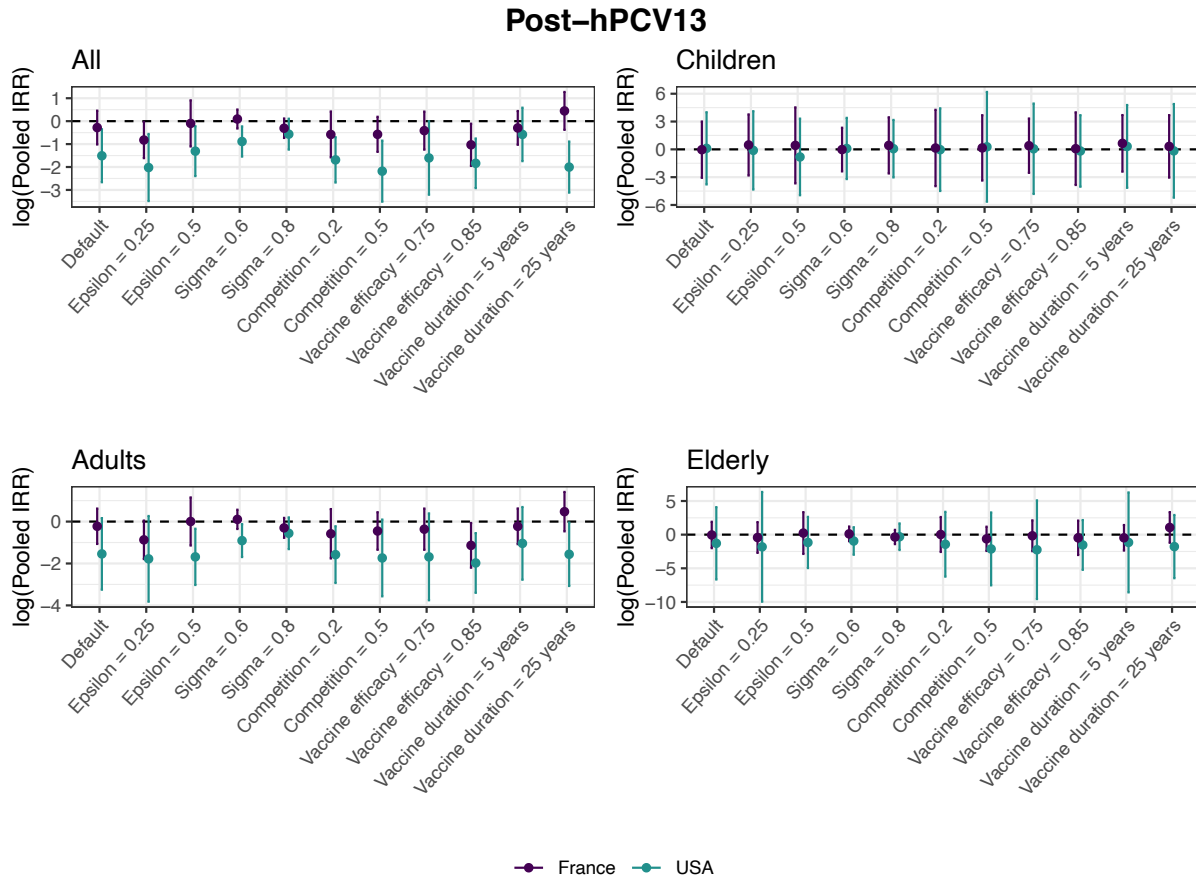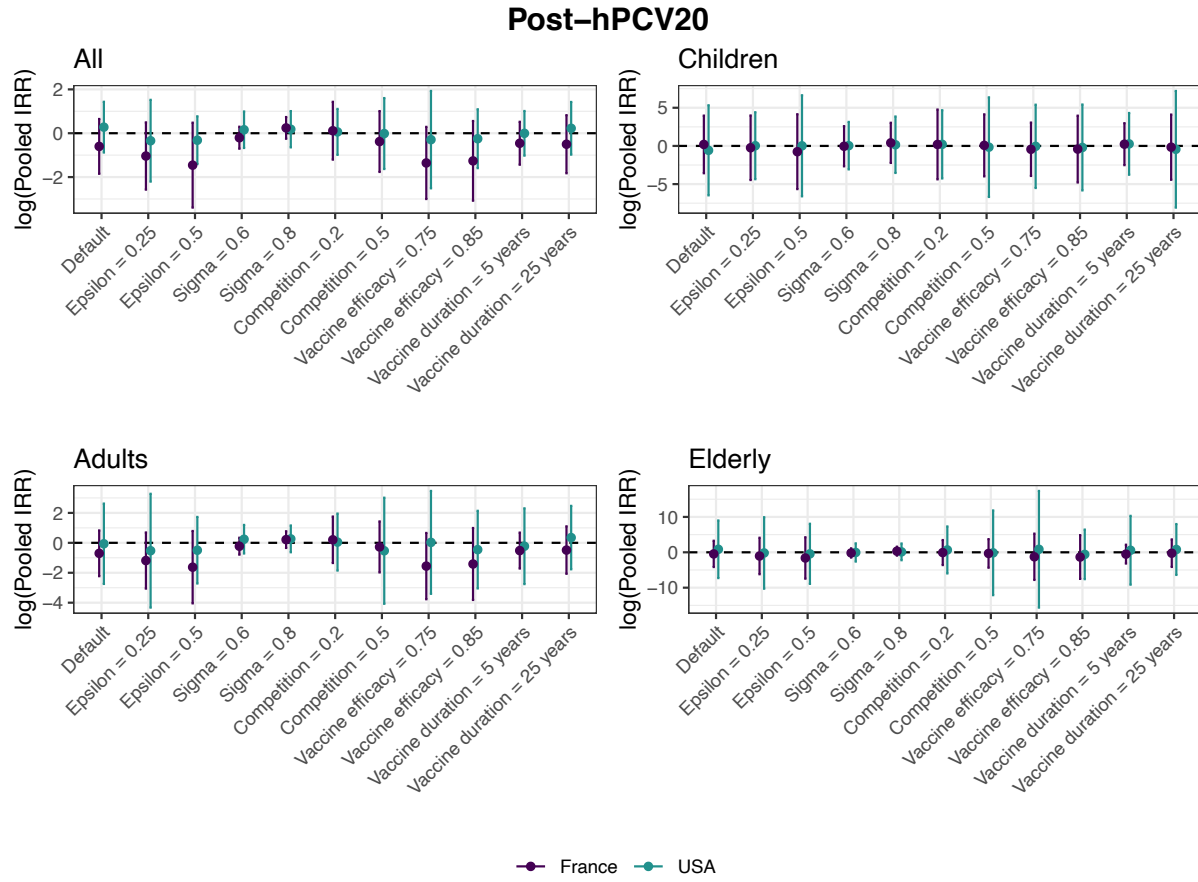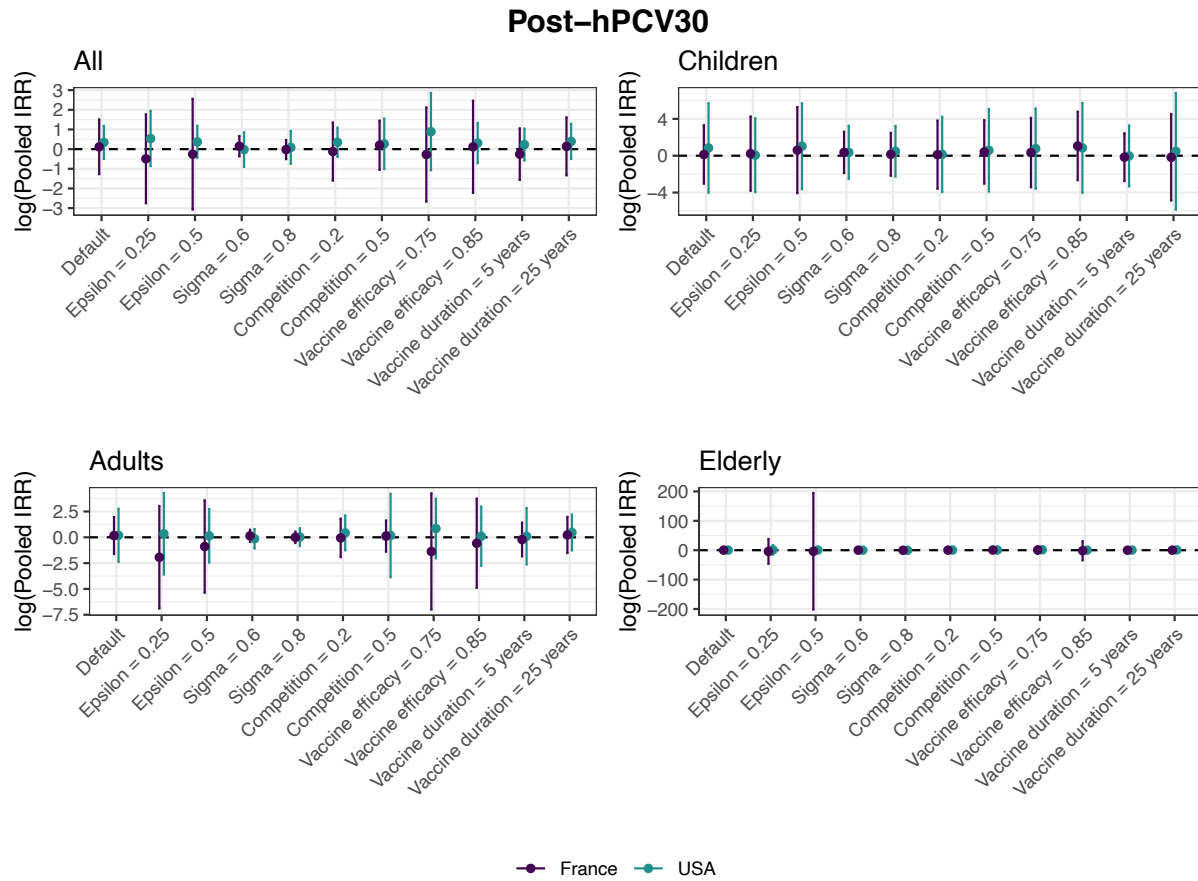

Figure 5.23: Comparison of IRR post-hPCV13 in French settings versus American setting across scenarios (varying each parameter respectively) under global invasiveness strategy and with the default parameter set. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.24: Comparison of IRR post-hPCV20 in French settings versus American setting across scenarios (varying each parameter respectively) under global invasiveness strategy and with the default parameter set. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

Figure 5.25: Comparison of IRR post-hPCV30 in French settings versus American setting across scenarios (varying each parameter respectively) under global invasiveness strategy and with the default parameter set. Default parameter set: $\varepsilon = 0.1, \sigma = 0.4, \theta = 0.1, v = 95\%$, $\omega = 10$ years. ε: non-specific immunity, σ: specific immunity, θ: competition, v: vaccine efficacy, ω: duration of vaccine protection.

5.4 Discussion

I developed an individual-based dynamic transmission model with 69 serotypes set in a population of 5000 individuals. Three vaccination coverage strategies were tested: serotype targeting based on (1) highest disease incidence, (2) highest local invasiveness, and (3) highest global invasiveness estimates. Vaccine valency was increased from hPCV13 to hPCV20 to hPCV30 every 10 years. Results across vaccination strategies were largely similar, as the serotypes that caused the most disease would be influenced by the invasiveness datasets. As such, the results showed that serotypes to be included in future PCV formulations based

on the French setting are serotypes 8, 10A, 10B, 12F, 16F, 24F, 25A, 33F, and 46, which were targeted both in the most highly invasive and most disease-causing simulations. This highlights that eventually the inclusion of the most disease-causing serotypes will be similar to the most highly invasive serotypes. The results were also quite consistent across sensitivity scenarios, demonstrated by the non-significant difference in both average carriage prevalence and pooled incidence rate ratios. These results would nonetheless benefit from further simulations or simulations with a much larger population size, as confidence intervals were wide in age groups with smaller populations. High end computational facilities would be required for such large scale simulations within populations of millions of people of the many serotypes and different vaccine types impacts.

Unfortunately, due to an error in the code, the US setting was not run under a local invasiveness or disease incidence strategy using its country-specific invasiveness dataset. Differences across countries in serotype replacement after implementation with PCV7 demonstrate that international implementation of a vaccine not specific to a setting may potentially be detrimental and less cost-effective. Future work should focus on expanding this model to other settings to compare the serotypes included between settings under the different vaccination strategies and with different invasiveness datasets. A global optimisation model would be complex but may ultimately elucidate whether a universal vaccine implementation would be beneficial on an international scale including developing countries. Despite the inability to compare between countries and strategies, this work predicts that vaccines would be most effective at reducing disease in the aggregate age group once the most invasive serotypes have been targeted. As cost effectiveness analyses are typically undertaken on a country-level, it would be reasonable to conclude that the hPCVs of increasing valency based on local invasiveness or disease would be cost-effective on an individual country level under the assumption of price parity (Chapter 2). Further work is needed to investigate the cost-effectiveness with respect to price differences between the vaccines.

While previous dynamic transmission models estimated the epidemiological impact of the currently licensed PCVs (Chapter 2), many grouped serotypes into VT and NVT and did not

take a serotype-specific approach applicable to vaccine design. Published models that took a serotype-specific approach did not model vaccine design. Only one previous study estimating the optimal vaccine types took a serotype-specific approach to model vaccine design using case-to-carrier ratios [356]. This study, set in the Finnish context, assumed a constant increase across NVTs as a group and did not account for transmission dynamics or serotype-specific durations of carriage that impact the recovery rates. Consequently, this study had limitations in its ability to address herd immunity and post-vaccination serotype replacement. Another study applied an ecological model to optimise pneumococcal vaccines using genomic data [331]. This study also did not model host transmission dynamics but showed that vaccine impact was sensitive to the circulating bacterial population and the integration of genomic data enabled the authors to incorporate pathogen evolution. While genomic data provides an exciting new avenue to model the transmission dynamics of each strain, the challenge rests with local infrastructure and resources particularly in LMICs. Serotype data requires less investment in expensive infrastructure. The model in this chapter could be applied to any setting with serotype data, uses serotype- and location-specific invasiveness data, is customisable and can be altered to change parameters across age groups and serotypes and incorporates host transmission dynamics that allow one to explore processes such as immunity, serotype replacement, herd immunity and vaccine impacts.

Nonetheless, the model presented here has many limitations. Models are only as good as their assumptions and the precision of the parameter estimates employed in simulations. The complexity of this model required many simplifying assumptions. For example, the model only considers carriage, and disease incidence is a snapshot estimated at the end of a vaccination period from invasiveness parameters, which do not change post-vaccination. Typically, only one serotype causes invasive disease, and therefore the disease incidence is likely over-estimated, as when the invasiveness is applied to serotype-specific carriage prevalence values, it does not take into account multi-serotype carriage. Furthermore, the invasiveness estimates are also over-estimated as they are estimated from one colony per carrier. As the values of disease incidence are static (only calculated once at the end of the

simulation period), the population cannot suffer or die from disease and therefore individuals cannot be removed from community transmission or from the population. The closed system may bias the system towards a better protected population, as individuals cannot die until the end of their life expectancy and therefore accumulate previous colonisations that will further protect them. This also means there are fewer 'new' susceptibles entering the population. This was reflected in the low overall pre-vaccination carriage prevalence values seen in the model. In reality, populations are highly dynamic with migration inward and outward, and deaths at all ages. This could be addressed by introducing demographic parameters such as age-specific death rates.

Due to the computational complexity of the model, point estimates and 95% credible intervals for invasiveness were used as opposed to drawing from respective distributions. Drawing from respective distributions would be a more statistically robust way to encompass uncertainty in these parameters. Additionally, point estimates of immunity parameters were also taken as those most reasonable according to Cobey and Lipsitch's paper [57]. Serotype-specific natural immunity in particular would be interesting to explore in a dynamic fashion, as it had a significant impact on the average carriage prevalence when elevated and this parameter was assumed not to change with increasing number of previous colonisations of a specific serotype. This later assumption may be an over-simplification. Furthermore, the number of events per time step was assumed to be fixed. Initially, an asynchronous time step model was set up where the time to the next event was estimated as the natural logarithm of a random number drawn from a uniform distribution divided by the sum of all rates. This resulted in significant slowing of the simulations, and hence an approximation for the number of events per time step was made. A fixed number of events per time step may not be appropriate as the number of infected individuals grow exponentially. More appropriately, the number of events for each time step could be drawn from a Poisson distribution with the rate parameter as the sum of all population rates. This too may result in a longer simulation time, but would be worth exploring further.

Many simplifying assumptions were made about colonisation. First, bacterial load and

superinfection (in which hosts can harbour multiple strains of a serotype and be counted as multiple carriers) were not considered to impact the FOI, as this is difficult to measure and implement. NP carriage and IPD have also been observed to change seasonally due to seasonal variations in viral coinfections [442], which were not accounted for in this model. Viral coinfections are thought to increase the shedding and therefore transmission rate of pneumococci [443]. Similarly, carriage duration in the model is independent of viral load from influenza, which has been known to influence recovery rates [419]. Furthermore, infectiousness occurred immediately after carriage acquisition, and did not change over the course of colonisation. Recovery rates depended only on previous colonisations and non-specific immunity, and did not change across populations, as they were based on the same infant data. These were used as a measure of serotype fitness [62], such that R_0 did not vary with transmission rates, which were kept constant across serotypes. Small differences between serotype-specific transmission rates were observed [289], however this, with serotype-specific competition rates, merits further exploration. Fitting carriage duration based on age rather than the number of previous colonisations may bias the intrinsic duration of carriage, since children tend to have several episodes during their first years of life [57]. However, due to the low pre-vaccination carriage prevalence values in children in the results presented here, it may be worth exploring age-specific immunity parameters in a future model such that immunity is dampened until individuals reach five years of age, or implement age-specific transmission rates to better reflect the reality where children have higher colonisation rates. Co-colonisation occurred in at least 10% of the cases in the infant refugee camp data [81], however the effect of increasing the maximum carrying capacity was not investigated as the number of co-colonisations becomes increasingly rare with additional multiple carriage. It is currently unclear whether vaccination has an impact on multi-serotype carriage [177, 444], though this would be an interesting avenue to explore in future iterations of this model. Moreover, the competition parameter in this work was kept constant across serotypes. A previous study found no differences in competition between most serotypes, although the authors noted a lack of serotype-specific data [427]. Competition may also change with age, something that was not accounted for here as a fixed parameter for competition was used.

Taken together, these simplifying assumptions highlight key areas where further studies are
needed to elucidate epidemiological parameters.

Pneumococcal populations are multi-serotype populations with niche and neutral mech-
anisms for coexistence that have been modelled previously [57, 239, 254]. A neutral-null
model is defined by Lipsitch et al as one with ecological neutrality, where the dynamics with
indistinguishable strains would depend only on ecological variables, and with population
genetic neutrality, where there is "no stable equilibrium frequency of the strains in the model"
[290]. In the case of the full serotype model however, there are clear violations of these
criteria that bias towards coexistence. First, serotypes have distinct ecological variables
based on the fitted data to inform recovery rates. While the effective contact rate is constant
across serotypes, the inclusion of serotype-specific immunity also violates the ecological
neutrality criterion as it alters the acquisition of a serotype by an individual. The population
genetic neutrality criterion does not seem to be violated as the simulations do not converge
to the same carriage prevalence values. Explicit mechanisms of coexistence are included in
this model. First, habitat heterogeneity, as the host population is divided into vaccinated
and unvaccinated individuals. Second, the addition of both non-specific and weak specific
immunity circumvents the competitive exclusion from the most fit serotype and increases
serotype diversity respectively. Exploring the parameter space between $\sigma = 0.4$ and 0.6
with concurrent changes in non-specific immunity would further explain the role that both
mechanisms of immunity play in the dynamics of carriage on serotype prevalence. Third,
competition exclusively impacts acquisition, and the clearance of any colonising serotype is
not affected by multiple carriage. Despite these violations of the criteria for a neutral-null
model, this work aimed to investigate the dynamics of serotype-targeting by vaccines and
therefore a neutral-null model would not be appropriate for this aim.

In conclusion, I have developed and investigated the predictions of an individual-based
stochastic model of pneumococcal serotype transmission and vaccine impact that elucidates
the difference between vaccine design strategies targeting serotypes based on global inva-
siveness, local invasiveness, and disease incidence. Although this model suffers from many

limitations, it illustrates the impacts of the vaccine design strategies adopted and broadly shows the impact of increasing serotype coverage based on these targeting strategies in high-income country settings. Ultimately, an invasiveness-targeting strategy alone may allow low invasiveness serotypes to proliferate unperturbed, although it would reduce the disease incidence significantly using the first thirteen valent formulation of the vaccine. This would only be an appropriate and ethical strategy if the low invasive but highly prevalent serotypes allowed to proliferate had low case fatality rates. While the disease incidence strategy that has been thus far followed will eventually include all the high-invasive serotypes as increasingly more serotypes are included in the vaccines and these are being more cheaply manufactured, a strategy targeting the locally invasive serotypes may also show significant reductions in disease.

5.5   Aim in this thesis

I have developed an individual-based dynamic transmission model that evaluates the impact of increasing the serotype coverage in PCVs based on three different vaccine strategies: (1) disease incidence, in which the most disease-causing serotypes are included, (2) local invasiveness, in which the serotypes with the highest local invasiveness are included, and (3) global invasiveness, in which the serotypes with the highest global invasiveness are included. Overall, despite non-significant results, all strategies reduce disease at similar levels across age groups and sensitivity scenarios, demonstrating that eventually the most critical serotypes will be included in all vaccine strategies. While there are areas requiring further improvements in the model and many assumptions around key parameters, it serves as a template for comparing vaccine targeting strategies based on different serotype compositions as inexpensive regional vaccines are licensed. It also serves as a good template for expanding the complexity of the model to run simulations in very large populations, with more complexity in parameter assignments for different serotypes and with the full demography of the human host population including age-specific death and birth rates. The model developed does also

serve to highlight which of the many key parameters that determine the epidemiology of

multi-serotype pneumococcal transmission require more focused study and measurement.

Chapter 6

Discussion and Concluding Remarks

This chapter summarises the research aims and findings of the thesis, provides critical limitations of the work presented and discusses future research in the field of pneumococcal vaccines.

6.1 Introduction

Streptococcus pneumoniae is a bacterial pathogen and commensal that is carried in the nasopharynx and is a significant cause of otitis media, pneumonia, and meningitis. Only seven of the approximately 100 serotypes circulating worldwide were initially included in PCV in 2000 before the number of serotypes included was expanded in subsequent years. While PCV reduced the disease incidence in regions of higher vaccine uptake, in part because of a herd immunity effect, a replacement effect was observed whereby carriage and disease were increasingly dominated by serotypes not included in the vaccine employed.

6.2 Summary of thesis aims and contributions

The aim of the research described in this thesis was to understand whether expanding valency of PCVs is desirable, given the balance of the benefits (direct protection and herd immunity) against the costs (financial and burden of serotype replacement) for past vaccines as well as what they might be for future vaccines. Additionally, this thesis aimed to explore serotype selection in expanded vaccines.

Dynamic transmission models can be employed to mimic the impact of serotype replacement and herd immunity effects to describe post-vaccination scenarios, whereas economic evaluations focused on costs and benefits can enable decision-makers to compare vaccines of increasing valency for implementation. In Chapter 2, I conducted a systematic literature review to examine published epidemiological and economic models with a focus on Europe and North America, and the assumptions made in these models for their potential contributions to future research, including parameter estimation, and immunisation policy. Twenty-nine dynamic transmission models and twenty-six economic models were identified and reviewed. Published models employed various structures, revealing several key uncertainties regarding the biology and epidemiology of pneumococcal infection. While models suggested that PCVs will reduce the burden of disease, the extent to which they are predicted to do so depended

on various assumptions regarding features of pneumococcal infection and epidemiology that strongly influence PCV cost-effectiveness conclusions. Such features include the duration of protection and competitive interactions between serotypes, both of which are unclear at present, but which directly relate to herd immunity and serotype replacement effects. In addition, few serotype-specific dynamic transmission models extend beyond two or three serotype grouping categories. Economic evaluations were not typically based on transmission dynamic models and hence omitted the indirect benefits arising from herd immunity in populations with sufficient vaccine coverage. The tools of transmission simulation and economic evaluation should be used in conjunction to inform decision-makers on vaccine implementation. To date, there have been few attempts to build economic evaluations on transmission dynamic models, and none on the situation prevailing in Europe and North America where vaccine uptake has in general been high.

In Chapter 3, I conducted a trend analysis assembling the available evidence for PCV impact on national IPD in European, North American and Australian settings. Significant effectiveness against VT IPD in infants was observed, although the impact on national IPD incidence varied by country due to differing patterns of serotype replacement. Currently, NVT serotypes 8, 9N, 15A and 23B are increasing in the countries assessed, although a variety of other NVTs are affecting each country and age group. Despite these common emerging serotypes, there has not been a dominant IPD serotype post-vaccination as there was pre-vaccination (serotype 14) or post-PCV7 (serotype 19A), suggesting that future vaccines with additional serotypes will be less effective at targeting and reducing IPD in global populations than previous PCVs. The rise of diverse NVTs in all settings' top-ranked IPD-causing serotypes emphasises the urgent need for surveillance data on serotype distribution and serotype-specific invasiveness post-vaccination to facilitate decision-making concerning both expanding current vaccination programmes and increasing vaccine valency. Based on the serotype trends of the countries included, universal vaccine expansion is not favourable due to diverging replacement, and results suggest that vaccines with different targeting mechanisms such as serotype agnostic vaccines or vaccines targeting local serotypes may be preferred. Diverging trends may be due

to pre-PCV bacterial population structures, environmental or sociological factors influencing transmission and replacement, antimicrobial use influencing antimicrobial resistance patterns, and variation in diagnostics and hospital practices [407]. All these factors may also influence serotype invasiveness.

Invasiveness, quantified as the hazard of carriage progressing to diseases, has been observed to be heterogeneous across pneumococci. Pneumococci are carried for varying durations, dependent on serotype and host age. Estimating invasiveness and understanding whether it is correlated strongly with serotype is crucial for designing effective vaccines. However, current epidemiological methods of estimating invasiveness have been confounded by post-vaccine changes in pneumococcal populations, necessitating the development of new methods to estimate this key parameter. In Chapter 4, a novel Bayesian framework was implemented to analyse epidemiological data from multiple locations to estimate both a dataset- and serotype-specific invasiveness (local invasiveness) and a single serotype-specific invasiveness across datasets (global invasiveness) for children less than 18 years and adults over 18 years. These estimates are absolute, allowing a direct comparison across datasets and between serotypes, which previous methods relying on relative measures, such as odds ratios, did not. Using Bayes' Factors, I was able to compare the two models to test whether serotype invasiveness varied by population. Model comparisons favour global invasiveness, which is unsurprising given that it is less complex as it estimates fewer parameters. However, serotype-specific local invasiveness distributions imply that there is geographic heterogeneity in serotypes' invasiveness, suggesting that serotypes may not have a universal invasiveness. There may also be population-independent factors that contribute to invasiveness, such as strain characteristics or serotype interactions. Pre-vaccination, high globally-invasive serotypes were 1, 4, 5 and 7F in children and 1, 4, 8 and 20 in adults although these estimates arise from a variable number and combination of datasets. The analysis suggests serotypes 8, 9N, 20 and 22F were among the most invasive serotypes globally not included in current vaccines. This framework will enable a strong evidence base to be built for future vaccine development and policy design, as it will be publicly available for others to expand upon with

their own datasets to re-estimate global invasiveness values. The analysis would particularly benefit from more datasets post-vaccination to determine whether invasiveness changes post-PCV implementation. Furthermore, a lack of consistent reporting on the number of swabs taken excluded many studies that would otherwise have been eligible for the invasiveness analysis. In the context of this thesis, invasiveness estimates were vital to translating modelled carriage prevalence to observed or predicted disease incidence in the dynamic transmission model.

In Chapter 5, I developed a multi-serotype individual-based dynamic transmission model (IBM) to evaluate the epidemiological impact of increasing serotype coverage in PCVs. IBMs are stochastic dynamic transmission models that simulate transmission based on a variety of epidemiological parameter estimates and behavioural rules governing transmission events. Using the local invasiveness estimates from Chapter 4, I applied this model to estimate the IPD incidence rate ratios in a French setting as a representative high-income European country, before and after vaccination with expanded valency based on three strategies. Because of the resulting complexity from the inclusion of 69 serotypes, I first investigated expanding serotype coverage to serotypes causing the most IPD, and then investigated expanding serotype coverage to serotypes with the highest local and global invasiveness respectively. Serotypes with both a high local and a high global invasiveness tended to be similar, although this may not apply to lower income countries. Though the hypothetical PCVs were all mostly successful in reducing VT carriage prevalence, implementing a PCV with increased serotype coverage based on local serotype invasiveness or circulation in carriage (and therefore disease) would be most effective in reducing IPD incidence on a country-level, as seen in practice. The serotypes that were added in the disease incidence strategy and the invasiveness strategies were very similar, and indicated that as PCVs are expanded the serotypes that cause the most disease and that are highly invasive will begin to overlap. Unfortunately, I was only able to compare multiple countries post-vaccination with the global invasiveness vaccine strategy. This strategy demonstrated more success in the US compared to France, as highly invasive serotypes in the US setting were more similar to highly invasive serotypes in the

global dataset compared to France. Despite only comparing one vaccination strategy across countries, this will be critical to understanding how differences in invasiveness shape the divergence in serotype landscape found in Chapter 3. It may be detrimental to develop and implement a PCV on a global level given stark differences between serotype landscapes and as a result of different serotype replacement patterns in different countries. The additional cost and complexity of developing such a vaccine should be evaluated against the potential benefits.

6.3   Discussion

The work presented in this thesis builds on previous work on serotype replacement surveillance, interpretation of serotype invasiveness, and dynamic pneumococcal transmission models, but it has a number of limitations largely relating to parameter estimation and data quality.

While Chapter 3 evaluated the serotype landscape in Europe, North America and Australian settings, LMICs constitute the majority of the burden of pneumococcal disease. Therefore, the global divergence in serotype replacement post-vaccination will be vital for vaccine impact epidemiology and will be increasingly important for vaccine manufacturing. Currently, serotype replacement in LMICs is not well-documented due to the dearth of robust disease surveillance systems [445]. As illustrated in Chapter 3, PCV impacts countries in a seemingly localised way, making a global transmission model highly ineffective at evaluating the impact of vaccination. Despite taking settings that are socioeconomically and demographically similar, and with similar pre-PCV serotype compositions in disease, I have shown that expanding a globally-licensed PCV may not be a sustainable intervention for targeting the pneumococcal disease burden as there is divergence in replacement even in similar settings.

In Chapter 4, results of the model comparison are by no means finalised because of the limited datasets included in the global invasiveness. The code is freely available such that anyone can

download it and upload their own studies to the datasets so that global invasiveness can be re-estimated and refined as more datasets are included [357]. The more datasets are included, the better we will understand whether invasiveness is truly a global phenomenon or whether it is subject to local specification as indicated by the results so far. Additionally, expanding on the limited datasets including both pre- and post-vaccination will hopefully provide some further clarification on whether invasiveness changes with mass vaccination. For example, such an analysis might clarify why similar serotype distributions were observed in the UK and the US in carriage but not in disease [323]. As previously mentioned, investigating a dataset-specific scale factor that may account for variations in local settings such that global invasiveness is adjusted for each location would be interesting, but this was beyond the scope of this thesis.

In order to keep the model in Chapter 5 as simple as possible, certain important hetero-geneities within and between populations were not included. Certain populations, such as Indigenous populations [105], or groups with comorbidities like chronic respiratory disease or immunocompromising conditions [446], are at higher risk of carriage and disease. Men are also at increased risk of IPD, even after vaccination [447, 448]. Pneumococcal pneumonia has a higher disease burden than IPD [445]. Here IPD as a disease outcome was not disaggregated. Differing vaccine efficacies to the individual serotypes included in PCVs, such as to serotype 3 in PCV13, or differences in dosing of vaccines was not included. Despite how crucial these multi-faceted risk factors are, these were unfortunately not able to be explored in this thesis, but would be interesting to investigate, as most of these have not been added as components to published dynamic transmission models thus far but can be easily added in this model. Further work on this chapter could also evaluate the effect of implementing a US-based vaccine in a French setting or vice versa, and to expand the analysis to other, less similar settings.

This work focused solely on the impact of vaccination on IPD. Although NIPD tends to follow the same patterns as IPD [236], NIPD is worth exploring independently as it is a key driver of vaccine cost-effectiveness. One reason for this is the frequency of AOM, often treated with

antibiotics, and therefore also exacerbating antimicrobial resistance [317]. As such, as vaccines prevent AOM and invasive disease, they also alleviate the burden of antimicrobial therapies on the healthcare system. Furthermore, while *S. pneumoniae* is responsible for a majority of severe bacterial pneumonia cases, vaccines are responsible for the marked decrease in bacterial pneumonia, resulting in pneumococcal pneumonia only accounting for around 7% of all severe pneumonia cases in children less than 5 years in a multi-site case-control study in seven low resource high burden settings [449]. This highlights the importance of NIPD surveillance, as another key measure of vaccine success.

More generally, infectious disease models are only as good as the data from which parameter estimates are derived and that form the quality of epidemiological and biological under- standing of what determines disease and transmission. The model that was developed in this thesis was unfortunately limited by the complexity (and therefore the computational requirements) of modelling each individual with over 60 serotypes, each of which may have different degrees of infectivity (β - the per capita rate at which individuals get infected with a defined serotype), which impacts the R_0. In the absence of other data, the model kept β constant, such that the R_0 varied only due to the serotype-specific carriage durations that were fitted to data. Although the model was complex and computationally expensive, it serves as a starting point for exploring the highly uncertain parameter space in the pneumococcal field.

Overall, key sources of uncertainty remain in the pneumococcal field of epidemiology. In- consistency in surveillance and reporting practices across countries, such as differences in age categories, made extrapolating information for comparison challenging. For example, the cut off age for children used in Chapter 3 (18 years, sometimes 16), Chapter 4 (18 years) and Chapter 5 (5 years) all differ due to the data available. More efforts should be made in surveillance to differentiate between infants and older children, and more generally in creat- ing consistency between surveillance practices and systems across countries as well as public availability. A crucial aspect of IPD surveillance is tracking emerging serotypes, but many surveillance systems such as the United States' Centers for Disease Control and Prevention publicly report aggregated IPD cases by location and age but not by serotype. Prior to the

coronavirus pandemic (COVID-19) when many surveillance systems were disrupted, the World Health Organization's (WHO) Global Invasive Bacterial Vaccine-Preventable Disease (IB-VPD) Surveillance Network reported that only 50 of its 194 member states reported data in August 2019 (Figure 6.1) [12]. These are countries that, at the lowest tier, report suspected meningitis in children less than five years old, and at the highest tier, have an active population-based surveillance. While almost half of African countries are included in this surveillance network, fewer than six have a national or sub-national reference laboratory, as these can be expensive to set up and operate. Although many European countries have pneumococcal surveillance in place and are involved in annual reporting and aggregation by the European Centre for Disease Control and Prevention (ECDC) [450], they are not included in the WHO's global effort. This highlights a lack of global coordination in identifying serotypes causing the greatest disease, despite the implementation of universal PCVs. It also emphasises the inequity between countries, as poorer countries may not have the resources to build expensive population-based surveillance systems. Recently, the The Pneumococcal Serotype Replacement and Distribution Estimation (PSERENADE) Project was commissioned by WHO to evaluate the impact of PCV10 and PCV13 on IPD incidence and serotype distribution [451]. The inclusion criteria for countries included age- and serotype-specific annual IPD case counts, at least one complete year of post-PCV data, at least 50% of isolates serotyped for at least one age stratum and at least 50% PCV uptake in at least one year post-PCV introduction, and no biases regarding testing or reporting. The project's results showed that 29 countries with PCV10/13 did not have surveillance in place, and 33 with PCV10/13 and IPD surveillance were not eligible for PSERENADE according to their criteria. Unfortunately, this project excludes countries that have not included PCV in their national immunisation schedules, such as China and India, and countries that do not use PCV10/13 but may use PCV7. Despite this, it will be a hugely beneficial step forward in the efforts to coordinate pneumococcal serotype- and age-specific IPD surveillance.

Another opportunity for pneumococcal research is expanding genomic surveillance, which has been ongoing since the 1990s but which has yet to be fully incorporated into interna-

tional surveillance, particularly in LMICs [452]. India, Nigeria, Democratic Republic of the Congo and Pakistan accounted for about 50% of the pneumococcal burden in 2015 but are only represented by 114, 67, 1, and 102 samples on the Global Pneumococcal Sequencing project database [453]. Beyond the UK, sampling in Europe has also been somewhat lacking. Genomic surveillance has been and will continue to be pivotal in surveillance of antimicrobial resistance, outbreaks, vaccine efficacy and elimination. Genomic surveillance is also important for detecting novel serotypes, something that will be critical to monitor as PCV valency is altered or expanded.

Parameter selection was another critical source of uncertainty. The modelling papers in Chapter 2 highlighted manufacturer bias that can easily be obscured in parameter selection when reporting guidelines are not used. Beyond this, there are a number of key epidemiological parameters that are still unknown and worth exploring in future research. First, because carriage prevalence alone does not explain invasive disease incidence [410], more work on transmission, competition or fitness parameters that encompass the interaction between serotypes would be beneficial in understanding why certain serotypes out-compete others and whether these would be good predictors of serotype growth. A key barrier in this work is that serotype-specific data can be noisy and lead to inaccurate predictions [410]. Previous work has shown that competition may be driven by capsule [306], and that between-serotype competition reduces acquisition [427], however the biological mechanism of competition is unknown and difficult to observe. Serotype-specific data disaggregated by age would be beneficial for understanding and modelling transmission and competition. So far, few efforts have been made to differentiate the transmission between age groups [436], and none have estimated the transmission to or from the elderly population in the community or household settings. Moreover, immunity parameters in dynamic transmission models are typically simplified. Both naturally-acquired and vaccine-induced immunity will not develop in all individuals homogeneously. In Chapter 5, the model did not incorporate within-host immunity that would be impacted by factors such as superinfection, bacterial load, or viral load, but this may affect the immune response as a certain threshold may alter the infectious period and is

necessary to trigger immunity [454]. Studies on key parameters from models involving herd immunity, serotype competition and the natural history of infection will be crucial to pneumococcal vaccine development and policy. These will also be crucial in determining the effect of future vaccines, both multi-valent and serotype-agnostic. Cross-functional discussions between industry and academia around parameters may help in setting a reasonable range for some parameters, such as cross-protection to non-vaccine serotypes, which impacted the results of the vaccine cost-effectiveness analyses. Such discussions would also organise and streamline efforts toward the most pressing questions pertaining to pneumococci.
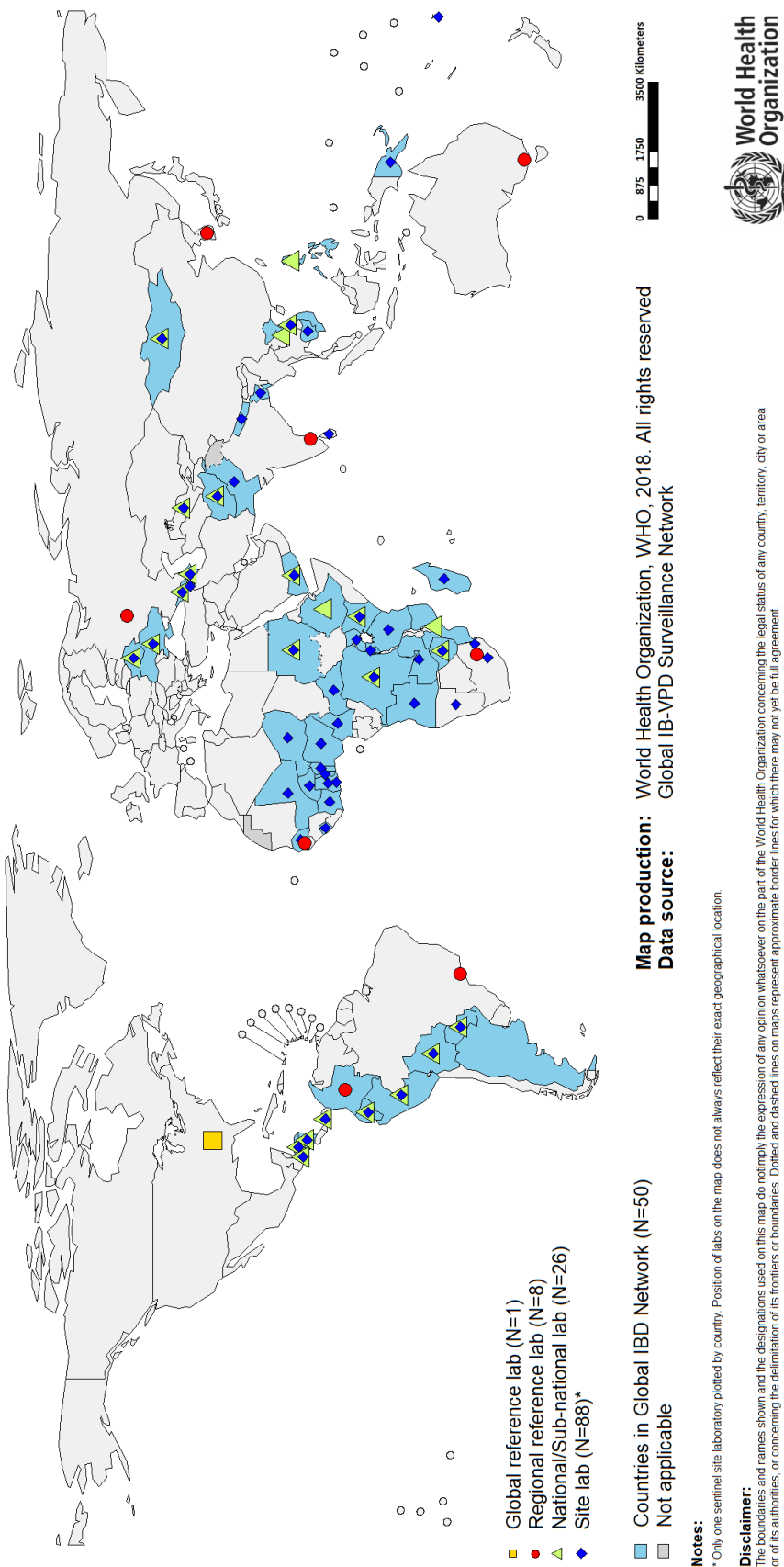
Figure 6.1: WHO Global Invasive Bacterial Vaccine-Preventable Disease (IB-VPD) Surveillance Network. Reproduced from [12].

Vaccines are generally cost-effective because of their impact on reducing the burden of disease, in turn reducing hospital visits, freeing up bed capacity and reducing the burden on caregivers [227]. However, vaccines are also the 'victims of their own success' [455], whereby once vaccines have successfully removed disease from a population, their costs are still visible [456]. The challenge with evaluating effectiveness and cost-effectiveness of expanded PCVs are their longer-term impacts, such as the burdens of emerging serotypes and antibiotic resistance which are difficult to predict. Tracking novel pneumococcal serotypes is important and ongoing, as the origin of these are sometimes homologous to other species, indicating uptake from exogenous species rather than previously undetected and/or low prevalence types [29, 457]. Vaccine costs are a critical determinant of cost-effectiveness, but these are usually vague estimates in cost-effectiveness analyses since tendering (i.e. vaccine bid) contracts are unknown, as are manufacturing and storage costs. In 2018, PCVs were reported to cost between US\$3-US\$132 per dose depending on the country [458], with ranges of costs overlapping between lower, middle and high income countries. The demand for pneumococcal vaccines is also expected to increase by more than 60 million doses over the next 10 years to 2030 with the addition of PCV in infant routine immunisation in 42 countries, including China and India [458]. Subsidised vaccines by GAVI and the Pan American Health Organization (PAHO) have enabled LMICs to access PCVs, whereas non-GAVI non-PAHO (i.e. self-procuring) middle income countries will hopefully overcome the affordability challenge they face with new manufacturers entering the market that will also meet the increasing global demand [458]. Assuming price parity across vaccines within a country, increasing serotype valency of PCVs would be cost-effective as more disease is averted at the same cost. However, one question that remains is whether it is more cost-effective to universally increase valency in a global vaccine, reducing manufacturing and marketing complexity for global vaccine manufacturers, or whether to locally increase valency on a country- or region-specific basis. Eventually, the end-products will be similar as the serotypes overlap between the two when a majority of serotypes are included, although it is unclear whether PCVs would still be equally effective against each serotype if it were expanded to this level of serotype coverage. A global optimisation comparing the disease cases averted

from increasing valency with local versus global circulating serotypes and varying vaccine efficacy for each serotype covered may elucidate this.

In conclusion, I sought to understand the impact of expanding serotype valency in PCVs in Europe and North America and found that expanding PCVs may not be a sustainable method for targeting pneumococcal disease in high income countries and therefore on a global level. Despite Europe and North America having similar demographics, the divergent trends in disease serotype replacement as well as differences in serotype-specific local invasiveness in these settings diminish the incentives for producing a "one size fits all" global serotype-targeting vaccine that is both biologically effective and cost-effective. This is particularly true as PCV-induced serotype replacement may have indirectly increased the antibiotic resistant population of pneumococci [459], further confounding the issue at stake, although this is somewhat controversial [460].

6.4   Future directions in pneumococcal vaccine development

This research work has demonstrated how vaccine design may be shaped by serotype invasiveness and the consequences of alternative targeting strategies. Ultimately, this work shows that targeting only the most invasive serotypes in future vaccines may ignore the prevalent serotypes that are causing disease but would still be effective at preventing disease over time. Targeting only the serotypes causing the most disease will eventually encompass the high invasive serotypes as the ability to cause disease hinges on the invasiveness. Based on the invasiveness and modelling chapters presented here, the serotypes critical for inclusion in PCVs are serotypes 8, 12F, and 33F. These serotypes were consistently included in both invasiveness-targeting strategies as well as the disease incidence-targeting strategy in the IBM developed here based on a French setting. These are included in the additional serotypes in the PCVs in development but not yet licensed: serotypes 8, 10A, 11A, 12F, 15B, 22F, 33F (Table 6.1). The other additional serotypes were frequently targeted by the vaccines modelled in Chapter 5 and highlighted in Chapter 3 and 4, except serotype 11A which was not included

in the global invasiveness strategy but has a high case-fatality rate [461] and is associated with antimicrobial resistance [462], providing validation that these additional serotypes are critical to target in the next formulations of PCV.

Alternative vaccines are being developed and tested to offset the effect of serotype replacement and the persistence of pneumococcal disease (Table 6.1). Current PCVs are expensive due to multi-step procedures involving many quality control steps as the product is heterologous. Limited licensing deals in other countries also leave countries susceptible to monopoly pricing, which has prompted in-house (i.e. within country) manufacturing and product innovation that will likely see an increase in pneumococcal vaccine suppliers in the coming years (Table 6.1) [463]. Lower cost PCVs are finally a reality, thanks in great part to key manufacturing optimisations in recombinant expression technology and conjugation chemistry that have made development less costly, as in the case of the Serum Institute of India's Pneumosil [464]. Lower cost PCVs present an opportunity, especially for LMICs, to create setting-specific vaccines to address diverging replacement trends. However, their mechanism is likely still to be carriage-blocking and therefore perpetuate the cycle of replacement in the nasopharynx with pathogens that may be more harmful. Other recent advances in production of conjugate vaccines, such as protein glycan coupling technology (PGCT) which does not require chemical conjugation and that uses *S. pneumoniae* carrier proteins, show promise for low cost and flexible vaccines that target disease but may also provide a degree of serotype-independent immunity [465]. These may be limited in which serotypes can be included due to the incompatibility between certain serotypes and coupling enzymes [466], as well as the varying degrees at which serotypes can be successfully expressed in the synthetic glycobiological development process [467].

Serotype-agnostic vaccines may be beneficial for avoiding perturbing the nasopharyngeal microbiota and the resulting replacement effect. The competition between VT and NVT would not alter the effectiveness of these vaccines, as they would block disease rather than carriage. Disease-blocking vaccines would require a higher vaccine coverage, since they would not benefit from herd immunity as carriage-blocking vaccines do but would theoretically still

offer the same if not better impact [480]. Due in part to the differences in demographics and life expectancy, the impact of PCVs on IPD cases averted differs between LMICs and HICs, as the pneumococcal burden is felt by different demographics [480]. The difference in impact of a disease-blocking vaccine between these settings would ultimately depend on the demographic being targeted for vaccination. Disease-blocking vaccines would supposedly not incur herd protection, and therefore routine childhood vaccination would not be as effective in HICs where the burden is primarily in adults and the elderly [480].

Furthermore, whether vaccination induces an increase [459, 481] or decrease [460] in frequencies of antibiotic resistance, a serotype-agnostic vaccine would help with a more ubiquitous targeting that would eliminate the need for antibiotics for pneumococcus and prevent vaccine-induced selection pressure that may elicit novel antigenic diversity via recombination of the capsule [482]. These benefits would only be reaped if the capsule-independent vaccines were not used in conjunction with PCVs, something that is still being tested [480]. Additionally, their success depends on vaccine targets being preserved in the pneumococcal population. Antibiotic resistance increases with antibiotic use in a population. Recently, a study by Davies et al. showed that the impact of serotype-independent vaccination on antibiotic resistance depended on the mechanism of the evolution of resistance and that the impact is greatest on the populations with the highest levels of antibiotic use and resistance [317]. As antimicrobial resistance is rising and threatens to be a potential pandemic in the coming years, sustained pneumococcal vaccination that reduces both the burden of invasive disease and the burden of antibiotic therapy will be critical.

6.5 Conclusions

While PCVs are very effective in reducing the burden of serious disease, and a number of lower cost vaccines will be entering the market in the next decade, the heterogeneity of serotype replacement and serotype invasiveness across countries risks reducing the benefit of universal PCVs of expanded valency. A number of exciting new technologies in vaccine

development and manufacturing will make serotype-independent or country-dependent targeting the future of preventing pneumococcal disease.

Table 6.1: Non-exhaustive list of PCVs in the research and development pipeline. Adapted from [4].

Vaccine (Manufacturer)	Serotypes	Phase	Notes	Reference
(Walvax)	13 (PCV13)	Approved	China-based, awaiting WHO pre-qualification	[468, 469]
20vPnC (Pfizer)	20 (PCV13 + 8, 10A, 11A, 12F, 15B, 22F, and 33F)	Phase 3	FDA-approval expected soon (2021)	[470]
V114 (Merck)	15 (PCV13 + 22F, 33F)	Phase 3	Licensure application submitted end of 2020	[471, 472]
(Beijing Minhai Biotechnology)	13 (PCV13)	Phase 3	China-based	[473]
SkyPneumo (SK Chemicals)	13	Phase 3	Patent dispute with Pfizer has delayed distribution	[4]
(Nanolek)	13	Phase 3	Russia-based	[463]
12vPHiD-CV (GSK)	12 (PCV10 + 6A, 19A)	Phase 2	Phase 1 recently completed (2020)	[474, 475]
(LG Chemicals)	N/A	Phase 2	South Korea-based	[474, 475]
(Panacea Biotech)	10 (PCV10)	Phase 2	India-based	[4, 476]
EuPCV15 (Eu-Biologics)	15 (PCV13 + 22F, 33F)	Phase 1	South Korea-based; mass-production of CRM17 carrier protein allows cheaper production	[477, 478]
(PnuVax)	13 (PCV13)	Pre-clinical	Canada-based	[479]

References

[1] C. Satzke, P. Turner, A. Virolainen-Julkunen, P. V. Adrian, M. Antonio, K. M. Hare, A. M. Henao-Restrepo, A. J. Leach, K. P. Klugman, B. D. Porter, R. Sa-Leao, J. A. Scott, H. Nohynek, K. L. O'Brien, and W. H. O. Pneumococcal Carriage Working Group. Standard method for detecting upper respiratory carriage of Streptococcus pneumoniae: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine*, 32(1):165–79, 2013. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2013.08.062. URL `https://www.ncbi.nlm.nih.gov/pubmed/24331112`.

[2] C. Satzke, E. M. Dunne, B. D. Porter, K. P. Klugman, E. K. Mulholland, and group PneuCarriage project. The PneuCarriage Project: A multi-centre comparative study to identify the best serotyping methods for examining pneumococcal carriage in vaccine evaluation studies. *PLoS Med*, 12(11):e1001903; discussion e1001903, 2015. ISSN 1549-1676 (Electronic) 1549-1277 (Linking). doi: 10.1371/journal.pmed.1001903. URL `https://www.ncbi.nlm.nih.gov/pubmed/26575033`.

[3] F. D. Schönbrodt and EJ Wagenmakers. Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25(1):128–142, 2018. ISSN 1531-5320. doi: 10.3758/s13423-017-1230-y. URL `https://doi.org/10.3758/s13423-017-1230-y`.

[4] UNICEF Supply Division. Pneumococcal conjugate vaccine: Supply and demand

update, 2020. URL `https://www.unicef.org/supply/media/4636/file/Pneumococcal-conjugate-vaccine-supply-update-July2020.pdf`.

[5] Control US Centers for Disease and Prevention. Pneumococcal photos: Photos of the disease and images of people affected by the disease, 2020. URL `https://www.cdc.gov/pneumococcal/about/photos.html`.

[6] N.J. Croucher, A. Løchen, and S.D. Bentley. Pneumococcal vaccines: Host interactions, population dynamics, and design principles. *Annual Review of Microbiology*, 72(1): 521–549, 2018. ISSN 0066-4227. doi: 10.1146/annurev-micro-090817-062338. URL `https://doi.org/10.1146/annurev-micro-090817-062338`.

[7] J. J. C. Drijkoningen and G. G. U. Rohde. Pneumococcal infection in adults: burden of disease. *Clinical Microbiology and Infection*, 20:45–51, 2014. ISSN 1198-743X. doi: https://doi.org/10.1111/1469-0691.12461. URL `http://www.sciencedirect.com/science/article/pii/S1198743X14601750`.

[8] B. Greenwood. The epidemiology of pneumococcal infection in children in the developing world. *Philos Trans R Soc Lond B Biol Sci*, 354(1384):777–85, 1999. doi: 10.1098/rstb.1999.0430. URL `https://www.ncbi.nlm.nih.gov/pubmed/10365403`.

[9] J.D. Cherry, G.J. Harrison, S.L. Kaplan, P.J. Hotez, and W.J. Steinbach. *Feigin and Cherry's Textbook of Pediatric Infectious Diseases*. Elsevier/Saunders, Philadelphia, PA, Eighth edition, 2019. ISBN 9781455711772 (two-volume set).

[10] H. L. Johnson, M. Deloria-Knoll, O. S. Levine, S. K. Stoszek, L. Freimanis Hance, R. Reithinger, L. R. Muenz, and K. L. O'Brien. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. *PLoS Med*, 7(10), 2010. ISSN 1549-1676 (Electronic) 1549-1277 (Linking). doi: 10.1371/journal.pmed.1000348. URL `https://www.ncbi.nlm.nih.gov/pubmed/20957191`.

[11] B. Wahl, K.L. O'Brien, A. Greenbaum, A. Majumder, L. Liu, Y. Chu, I. Lukšić, H. Nair, D.A. McAllister, H. Campbell, I. Rudan, R. Black, and M.D. Knoll. Burden of Strep-

tococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *The Lancet Global Health*, 6(7):e744–e757, 2018. ISSN 2214-109X. doi: 10.1016/S2214-109X(18) 30247-X. URL https://pubmed.ncbi.nlm.nih.gov/29903376.

[12] WHO Global Invasive Bacterial Vaccine-Preventable Disease and Rotavirus and Pediatric Diarrhea Surveillance Networks Bulletin, 2019. URL https://mailchi.mp/b217dc802dcc/who-ib-vpd-and-rotavirus-surveillance-bulletin-june-1910809?e=225412eff9.

[13] L. Pasteur. Note sur la maladie nourvelle provoquee par la salive d'un enfant mort de la rage. *Bulletin de l'Academie de Medicine*, 10:94–103, 1881.

[14] G. M. Sternberg. A fatal form of septicaemia in the rabbit, produced by the subcutaneous injection of human saliva: an experimental research. *Natl Board of Health Bulletin*, 2:781–783, 1881.

[15] E. Klebs. Beiträge zur Kenntniss der pathogenen Schistomyceten. *Archiv für experimentelle Pathologie und Pharmakologie*, 5(4):350–377, 1876. ISSN 1432-1912. doi: 10.1007/BF01976922. URL https://doi.org/10.1007/BF01976922.

[16] R. Koch. *Zur Untersuchung von pathogenen Organismen*. Robert Koch-Institut, 1881. doi: http://dx.doi.org/10.25646/5071.

[17] C. J. Eberth. Zur Kenntniss der mykotischen Prozesse. *Deutsches Archiv für klinische Medicin*, 28:1–42, 1881.

[18] US Centers for Disease Control and Prevention. Risk factors and transmission, 2021. URL https://www.cdc.gov/pneumococcal/about/risk-transmission.html.

[19] B. M. Gray, 3rd Converse, G. M., and Jr. Dillon, H. C. Epidemiologic studies of Streptococcus pneumoniae in infants: acquisition, carriage, and infection during the

first 24 months of life. *J Infect Dis*, 142(6):923–33, 1980. URL `https://www.ncbi.nlm.nih.gov/pubmed/7462701`.

[20] T. M. File. Community-acquired pneumonia. *Lancet*, 362(9400):1991–2001, 2003. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(03)15021-0. URL `https://www.ncbi.nlm.nih.gov/pubmed/14683661`.

[21] W. S. Lim, S. V. Baudouin, R. C. George, A. T. Hill, C. Jamieson, I. Le Jeune, J. T. Macfarlane, R. C. Read, H. J. Roberts, M. L. Levy, M. Wani, M. A. Woodhead, and B. T. S. Standards of Care Committee Pneumonia Guidelines Committee of the. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*, 64 Suppl 3:iii1–55, 2009. ISSN 1468-3296 (Electronic) 0040-6376 (Linking). doi: 10.1136/thx.2009.121434. URL `https://www.ncbi.nlm.nih.gov/pubmed/19783532`.

[22] A. Ruiz-Gonzalez, M. Falguera, A. Nogues, and M. Rubio-Caballero. Is Streptococcus pneumoniae the leading cause of pneumonia of unknown etiology? A microbiologic study of lung aspirates in consecutive patients with community-acquired pneumonia. *Am J Med*, 106(4):385–90, 1999. doi: 10.1016/s0002-9343(99)00050-9. URL `https://www.ncbi.nlm.nih.gov/pubmed/10225239`.

[23] D. R. Cundell, N. P. Gerard, C. Gerard, I. Idanpaan-Heikkila, and E. I. Tuomanen. Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet-activating factor. *Nature*, 377(6548):435–8, 1995. doi: 10.1038/377435a0. URL `https://www.ncbi.nlm.nih.gov/pubmed/7566121`.

[24] N. J. Croucher, J. J. Campo, T. Q. Le, X. Liang, S. D. Bentley, W. P. Hanage, and M. Lipsitch. Diverse evolutionary patterns of pneumococcal antigens identified by pangenome-wide immunological screening. *Proc Natl Acad Sci U S A*, 114(3):E357–E366, 2017. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1613937114. URL `https://www.ncbi.nlm.nih.gov/pubmed/28053228`.

[25] C. Hyams, E. Camberlein, J. M. Cohen, K. Bax, and J. S. Brown. The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by

multiple mechanisms. *Infect Immun*, 78(2):704–15, 2010. ISSN 1098-5522 (Electronic) 0019-9567 (Linking). doi: 10.1128/IAI.00881-09. URL `https://www.ncbi.nlm.nih.gov/pubmed/19948837`.

[26] U. B. Skov Sorensen, J. Blom, A. Birch-Andersen, and J. Henrichsen. Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in pneumococci. *Infect Immun*, 56(8):1890–6, 1988. URL `https://www.ncbi.nlm.nih.gov/pubmed/3397179`.

[27] O. T. Avery and R. Dubos. The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med*, 54(1):73–89, 1931. URL `https://www.ncbi.nlm.nih.gov/pubmed/19869903`.

[28] K. A. Geno, G. L. Gilbert, J. Y. Song, I. C. Skovsted, K. P. Klugman, C. Jones, H. B. Konradsen, and M. H. Nahm. Pneumococcal capsules and their types: Past, present, and future. *Clin Microbiol Rev*, 28(3):871–99, 2015. ISSN 1098-6618 (Electronic) 0893-8512 (Linking). doi: 10.1128/CMR.00024-15. URL `https://www.ncbi.nlm.nih.gov/pubmed/26085553`.

[29] F. Ganaie, J. S. Saad, L. McGee, A. J. van Tonder, S. D. Bentley, S. W. Lo, R. A. Gladstone, P. Turner, J. D. Keenan, R. F. Breiman, and M. H. Nahm. A new pneumococcal capsule type, 10D, is the 100th serotype and has a large *cps* fragment from an oral streptococcus. *mBio*, 11(3):e00937–20, 2020. doi: 10.1128/mBio.00937-20. URL `https://mbio.asm.org/content/mbio/11/3/e00937-20.full.pdf`.

[30] C.S. Angel, M. Ruzek, and M.K. Hostetter. Degradation of C3 by Streptococcus pneumoniae. *The Journal of Infectious Diseases*, 170(3):600–608, 1994. ISSN 0022-1899. doi: 10.1093/infdis/170.3.600. URL `https://doi.org/10.1093/infdis/170.3.600`.

[31] M. K. Hostetter. Serotypic variations among virulent pneumococci in deposition and degradation of covalently bound C3b: Implications for phagocytosis and antibody production. *The Journal of Infectious Diseases*, 153(4):682–693, 1986. ISSN 0022-1899. doi: 10.1093/infdis/153.4.682. URL `https://doi.org/10.1093/infdis/153.4.682`.

[32] E. Alonso De Velasco, A. F. Verheul, J. Verhoef, and H. Snippe. Streptococcus pneumoniae: virulence factors, pathogenesis, and vaccines. *Microbiological Reviews*, 59 (4):591–603, 1995. ISSN 0146-0749. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC239389/`.

[33] D. A. Watson and D. M. Musher. Interruption of capsule production in Streptococcus pneumoniae serotype 3 by insertion of transposon Tn916. *Infect Immun*, 58(9):3135–8, 1990. URL `https://www.ncbi.nlm.nih.gov/pubmed/2167295`.

[34] L. E. Keller, D. A. Robinson, and L. S. McDaniel. Nonencapsulated Streptococcus pneumoniae: Emergence and pathogenesis. *MBio*, 7(2):e01792, 2016. doi: 10.1128/mBio. 01792-15.

[35] M. J. Jedrzejas. Pneumococcal virulence factors: Structure and function. *Microbiology and Molecular Biology Reviews*, 65(2):187–207, 2001. ISSN 1092-2172 1098-5557. doi: 10.1128/MMBR.65.2.187-207.2001. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC99024/`.

[36] L. Richards, D. M. Ferreira, E. N. Miyaji, P. W. Andrew, and A. Kadioglu. The immunising effect of pneumococcal nasopharyngeal colonisation; protection against future colonisation and fatal invasive disease. *Immunobiology*, 215(4):251–63, 2010. ISSN 1878-3279 (Electronic) 0171-2985 (Linking). doi: 10.1016/j.imbio.2009.12.004. URL `https://www.ncbi.nlm.nih.gov/pubmed/20071053`.

[37] R. A. Lock, D. Hansman, and J. C. Paton. Comparative efficacy of autolysin and pneumolysin as immunogens protecting mice against infection by Streptococcus pneumoniae. *Microb Pathog*, 12(2):137–43, 1992. URL `https://www.ncbi.nlm.nih.gov/pubmed/1350045`.

[38] M. A. Zafar, Y. Wang, S. Hamaguchi, and J. N. Weiser. Host-to-host transmission of Streptococcus pneumoniae is driven by its inflammatory toxin, pneumolysin. *Cell Host & Microbe*, 21(1):73–83, 2017. ISSN 1934-6069 1931-3128. doi: 10.1016/j.

chom.2016.12.005. URL `https://pubmed.ncbi.nlm.nih.gov/28081446https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5267320/`.

[39] Muna F. Abry, K.M. Kimenyi, F.O. Osowo, W.O. Odhiambo, S.O. Sewe, and Benard W. Kulohoma. Genetic diversity of the pneumococcal CbpA: Implications for next-generation vaccine development. *Human Vaccines and Immunotherapeutics*, 11(5): 1261–1267, 2015. ISSN 2164-5515 2164-554X. doi: 10.1080/21645515.2015.1021521. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4514174/`.

[40] A. M. Berry and J. C. Paton. Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of Streptococcus pneumoniae. *Infect Immun*, 64(12): 5255–62, 1996.

[41] A. Dintilhac, G. Alloing, C. Granadel, and J. P. Claverys. Competence and virulence of Streptococcus pneumoniae: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. *Mol Microbiol*, 25(4): 727–39, 1997.

[42] L. J. McAllister, H. J. Tseng, A. D. Ogunniyi, M. P. Jennings, A. G. McEwan, and J. C. Paton. Molecular analysis of the psa permease complex of Streptococcus pneumoniae. *Mol Microbiol*, 53(3):889–901, 2004. doi: 10.1111/j.1365-2958.2004.04164.x. URL `https://www.ncbi.nlm.nih.gov/pubmed/15255900`.

[43] H. J. Tseng, A. G. McEwan, J. C. Paton, and M. P. Jennings. Virulence of Streptococcus pneumoniae: PsaA mutants are hypersensitive to oxidative stress. *Infect Immun*, 70(3): 1635–9, 2002. URL `https://www.ncbi.nlm.nih.gov/pubmed/11854257`.

[44] M. J. Jedrzejas, E. Lamani, and R. S. Becker. Characterization of selected strains of pneumococcal surface protein A. *J Biol Chem*, 276(35):33121–8, 2001. doi: 10.1074/jbc. M103304200. URL `https://www.ncbi.nlm.nih.gov/pubmed/11413137`.

[45] D. M. Ferreira, D. R. Neill, M. Bangert, J. F. Gritzfeld, N. Green, A. K. Wright, S. H. Pennington, L. Bricio-Moreno, A. T. Moreno, E. N. Miyaji, A. D. Wright, A. M.

Collins, D. Goldblatt, A. Kadioglu, and S. B. Gordon. Controlled human infection and rechallenge with Streptococcus pneumoniae reveals the protective efficacy of carriage in healthy adults. *Am J Respir Crit Care Med*, 187(8):855–64, 2013. ISSN 1535-4970 (Electronic) 1073-449X (Linking). doi: 10.1164/rccm.201212-2277OC. URL https://www.ncbi.nlm.nih.gov/pubmed/23370916.

[46] D. M. Weinberger, R. Dagan, N. Givon-Lavi, G. Regev-Yochay, R. Malley, and M. Lipsitch. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. *J Infect Dis*, 197(11):1511–8, 2008. doi: 10.1086/587941. URL https://www.ncbi.nlm.nih.gov/pubmed/18471062.

[47] P. Klein Klouwenberg and L. Bont. Neonatal and infantile immune responses to encapsulated bacteria and conjugate vaccines. *Clinical and Developmental Immunology*, 2008:628963, 2008. ISSN 1740-2522 1740-2530. doi: 10.1155/2008/628963. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2553187/.

[48] G. Blanchard-Rohner and A. J. Pollard. Long-term protection after immunization with protein-polysaccharide conjugate vaccines in infancy. *Expert Rev Vaccines*, 10(5):673–84, 2011. ISSN 1744-8395 (Electronic) 1476-0584 (Linking). doi: 10.1586/erv.11.14. URL https://www.ncbi.nlm.nih.gov/pubmed/21604987.

[49] M. Lipsitch, C. G. Whitney, E. Zell, T. Kaijalainen, R. Dagan, and R. Malley. Are anticapsular antibodies the primary mechanism of protection against invasive pneumococcal disease? *PLoS Med*, 2(1):e15, 2005. doi: 10.1371/journal.pmed.0020015. URL https://doi.org/10.1371/journal.pmed.0020015.

[50] S. P. Jochems, J. N. Weiser, R. Malley, and D. M. Ferreira. The immunological mechanisms that control pneumococcal carriage. *PLOS Pathogens*, 13(12):e1006665, 2017. doi: 10.1371/journal.ppat.1006665. URL https://doi.org/10.1371/journal.ppat.1006665.

[51] M. N. Khan and M. E. Pichichero. The host immune dynamics of pneumococcal colonization: implications for novel vaccine development. *Hum Vaccin Immunother*, 10(12):

3688–99, 2014. ISSN 2164-554X (Electronic) 2164-5515 (Linking). doi: 10.4161/21645515. 2014.979631. URL `https://www.ncbi.nlm.nih.gov/pubmed/25668673`.

[52] F. D. Batista and N. E. Harwood. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol*, 9(1):15–27, 2009. ISSN 1474-1741 (Electronic) 1474-1733 (Linking). doi: 10.1038/nri2454. URL `https://www.ncbi.nlm.nih.gov/pubmed/19079135`.

[53] A. Cerutti. The regulation of IgA class switching. *Nat Rev Immunol*, 8(6):421–34, 2008. doi: 10.1038/nri2322. URL `https://www.ncbi.nlm.nih.gov/pubmed/18483500`.

[54] D. Goldblatt, M. Hussain, N. Andrews, L. Ashton, C. Virta, A. Melegaro, R. Pebody, R. George, A. Soininen, J. Edmunds, N. Gay, H. Kayhty, and E. Miller. Antibody responses to nasopharyngeal carriage of Streptococcus pneumoniae in adults: a longitudinal household study. *J Infect Dis*, 192(3):387–93, 2005. doi: 10.1086/431524.

[55] P. C. Hill, Y. B. Cheung, A. Akisanya, K. Sankareh, G. Lahai, B. M. Greenwood, and R. A. Adegbola. Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian infants: a longitudinal study. *Clin Infect Dis*, 46(6):807–14, 2008. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1086/528688. URL `https://www.ncbi.nlm.nih.gov/pubmed/18279039`.

[56] S. M. Granat, J. Ollgren, E. Herva, Z. Mia, K. Auranen, and P. H. Makela. Epidemiological evidence for serotype-independent acquired immunity to pneumococcal carriage. *J Infect Dis*, 200(1):99–106, 2009. doi: 10.1086/599364. URL `https://www.ncbi.nlm.nih.gov/pubmed/19469705`.

[57] S. Cobey and M. Lipsitch. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science*, 335(6074):1376–80, 2012. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1215947. URL `https://www.ncbi.nlm.nih.gov/pubmed/22383809`.

[58] M. E. Conley, A. K. Dobbs, D. M. Farmer, S. Kilic, K. Paris, S. Grigoriadou, E. Coustan-Smith, V. Howard, and D. Campana. Primary B cell immunodeficiencies: comparisons and contrasts. *Annu Rev Immunol*, 27:199–227, 2009. doi: 10.1146/annurev.immunol. 021908.132649. URL `https://www.ncbi.nlm.nih.gov/pubmed/19302039`.

[59] C. Picard, J. L. Casanova, and A. Puel. Infectious diseases in patients with IRAK-4, MyD88, NEMO, or IkappaBalpha deficiency. *Clin Microbiol Rev*, 24(3):490–7, 2011. ISSN 1098-6618 (Electronic) 0893-8512 (Linking). doi: 10.1128/CMR.00001-11. URL `https://www.ncbi.nlm.nih.gov/pubmed/21734245`.

[60] R. M. Douglas, J. C. Paton, S. J. Duncan, and D. J. Hansman. Antibody response to pneumococcal vaccination in children younger than five years of age. *J Infect Dis*, 148 (1):131–7, 1983. URL `https://www.ncbi.nlm.nih.gov/pubmed/6886479`.

[61] M. W. Mureithi, A. Finn, M. O. Ota, Q. Zhang, V. Davenport, T. J. Mitchell, N. A. Williams, R. A. Adegbola, and R. S. Heyderman. T cell memory response to pneumococcal protein antigens in an area of high pneumococcal carriage and disease. *J Infect Dis*, 200(5):783–93, 2009. doi: 10.1086/605023. URL `https://www.ncbi.nlm.nih.gov/pubmed/19642930`.

[62] M. Lipsitch, O. Abdullahi, A. D'Amour, W. Xie, D. M. Weinberger, E. Tchetgen Tchetgen, and J. A. Scott. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology*, 23(4):510–9, 2012. ISSN 1531-5487 (Electronic) 1044-3983 (Linking). doi: 10.1097/EDE.0b013e31824f2f32. URL `https://www.ncbi.nlm.nih.gov/pubmed/22441543`.

[63] L. Hogberg, P. Geli, H. Ringberg, E. Melander, M. Lipsitch, and K. Ekdahl. Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. *J Clin Microbiol*, 45(3):948–52, 2007. doi: 10.1128/JCM. 01913-06. URL `https://www.ncbi.nlm.nih.gov/pubmed/17202280`.

[64] J. A. Scott, A. J. Hall, R. Dagan, J. M. Dixon, S. J. Eykyn, A. Fenoll, M. Hortal, L. P. Jette,

J. H. Jorgensen, F. Lamothe, C. Latorre, J. T. Macfarlane, D. M. Shlaes, L. E. Smart, and A. Taunay. Serogroup-specific epidemiology of Streptococcus pneumoniae: associations with age, sex, and geography in 7,000 episodes of invasive disease. *Clin Infect Dis*, 22 (6):973–81, 1996.

[65] A. B. Brueggemann, D. T. Griffiths, E. Meats, T. Peto, D. W. Crook, and B. G. Spratt. Clonal relationships between invasive and carriage Streptococcus pneumoniae and serotype- and clone-specific differences in invasive disease potential. *J Infect Dis*, 187 (9):1424–32, 2003. doi: 10.1086/374624. URL https://www.ncbi.nlm.nih.gov/pubmed/12717624.

[66] A. B. Brueggemann, T. E. Peto, D. W. Crook, J. C. Butler, K. G. Kristinsson, and B. G. Spratt. Temporal and geographic stability of the serogroup-specific invasive disease potential of Streptococcus pneumoniae in children. *J Infect Dis*, 190(7):1203–11, 2004. doi: 10.1086/423820. URL https://www.ncbi.nlm.nih.gov/pubmed/15346329.

[67] D. Bogaert, R. De Groot, and P. W. Hermans. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *Lancet Infect Dis*, 4(3):144–54, 2004. doi: 10.1016/S1473-3099(04)00938-7. URL https://www.ncbi.nlm.nih.gov/pubmed/14998500.

[68] K. L. Sleeman, D. Griffiths, F. Shackley, L. Diggle, S. Gupta, M. C. Maiden, E. R. Moxon, D. W. Crook, and T. E. Peto. Capsular serotype-specific attack rates and duration of carriage of Streptococcus pneumoniae in a population of children. *J Infect Dis*, 194(5): 682–8, 2006. doi: 10.1086/505710.

[69] W. P. Hausdorff, D. R. Feikin, and K. P. Klugman. Epidemiological differences among pneumococcal serotypes. *Lancet Infect Dis*, 5(2):83–93, 2005. ISSN 1473-3099. doi: https://doi.org/10.1016/S1473-3099(05)01280-6. URL http://www.sciencedirect.com/science/article/pii/S1473309905012806.

[70] B. Simell, K. Auranen, H. Kayhty, D. Goldblatt, R. Dagan, K. L. O'Brien, and Group Pneumococcal Carriage. The fundamental link between pneumococcal carriage and

disease. *Expert Rev Vaccines*, 11(7):841–55, 2012. ISSN 1744-8395 (Electronic) 1476-0584 (Linking). doi: 10.1586/erv.12.53. URL https://www.ncbi.nlm.nih.gov/pubmed/22913260.

[71] B. D. Gessner, J. E. Mueller, and S. Yaro. African meningitis belt pneumococcal disease epidemiology indicates a need for an effective serotype 1 containing vaccine, including for older children and adults. *BMC Infect Dis*, 10:22, 2010. ISSN 1471-2334 (Electronic) 1471-2334 (Linking). doi: 10.1186/1471-2334-10-22. URL https://www.ncbi.nlm.nih.gov/pubmed/20146815.

[72] Y. Traore, T. A. Tameklo, B. M. Njanpop-Lafourcade, M. Lourd, S. Yaro, D. Niamba, A. Drabo, J. E. Mueller, J. L. Koeck, and B. D. Gessner. Incidence, seasonality, age distribution, and mortality of pneumococcal meningitis in Burkina Faso and Togo. *Clin Infect Dis*, 48 Suppl 2:S181–9, 2009. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1086/596498. URL https://www.ncbi.nlm.nih.gov/pubmed/19191614.

[73] M. M. Rovers, A. G. Schilder, G. A. Zielhuis, and R. M. Rosenfeld. Otitis media. *Lancet*, 363(9407):465–73, 2004. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(04)15495-0. URL https://www.ncbi.nlm.nih.gov/pubmed/14962529.

[74] K. L. Sutton, H. T. Banks, and C. Castillo-Chavez. Estimation of invasive pneumococcal disease dynamics parameters and the impact of conjugate vaccination in Australia. *Math Biosci Eng*, 5(1):175–204, 2008. URL https://www.ncbi.nlm.nih.gov/pubmed/18193937.

[75] A. C. Berical, D. Harris, C. S. Dela Cruz, and J. D. Possick. Pneumococcal vaccination strategies. An update and perspective. *Ann Am Thorac Soc*, 13(6):933–44, 2016. ISSN 2325-6621 (Electronic) 2325-6621 (Linking). doi: 10.1513/AnnalsATS.201511-778FR. URL https://www.ncbi.nlm.nih.gov/pubmed/27088424.

[76] A. Sinclair, X. Xie, M. Teltscher, and N. Dendukuri. Systematic review and meta-analysis of a urine-based pneumococcal antigen test for diagnosis of community-acquired

pneumonia caused by Streptococcus pneumoniae. *Journal of Clinical Microbiology*, 51(7):2303–2310, 2013. doi: 10.1128/jcm.00137-13. URL `https://jcm.asm.org/content/jcm/51/7/2303.full.pdf`.

[77] P. H. Edelstein, C. S. Jørgensen, and L. A. Wolf. Performance of the ImmuView and BinaxNOW assays for the detection of urine and cerebrospinal fluid Streptococcus pneumoniae and Legionella pneumophila serogroup 1 antigen in patients with Legionnaires' disease or pneumococcal pneumonia and meningitis. *PLOS ONE*, 15(8): e0238479, 2020. doi: 10.1371/journal.pone.0238479. URL `https://doi.org/10.1371/journal.pone.0238479`.

[78] Scott F. Dowell, Robert L. Garman, Gang Liu, Orin S. Levine, and Yong-Hong Yang. Evaluation of Binax NOW, an assay for the detection of pneumococcal antigen in urine samples, performed among pediatric patients. *Clinical Infectious Diseases*, 32(5):824–825, 2001. ISSN 1058-4838. doi: 10.1086/319205. URL `https://doi.org/10.1086/319205`.

[79] L. B. Reller, M. P. Weinstein, A. M. Werno, and D. R. Murdoch. Laboratory diagnosis of invasive pneumococcal disease. *Clinical Infectious Diseases*, 46(6):926–932, 2008. ISSN 1058-4838. doi: 10.1086/528798. URL `https://doi.org/10.1086/528798`.

[80] A.J. Blaschke. Interpreting assays for the detection of Streptococcus pneumoniae. *Clinical Infectious Diseases*, 52 Suppl 4(Suppl 4):S331–S337, 2011. ISSN 1537-6591 1058-4838. doi: 10.1093/cid/cir048. URL `https://pubmed.ncbi.nlm.nih.gov/21460292`.

[81] P. Turner, J. Hinds, C. Turner, A. Jankhot, K. Gould, S. D. Bentley, F. Nosten, and D. Goldblatt. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *J Clin Microbiol*, 49(5):1784–9, 2011. ISSN 1098-660X (Electronic) 0095-1137 (Linking). doi: 10.1128/JCM.00157-11. URL `https://www.ncbi.nlm.nih.gov/pubmed/21411589`.

[82] R. E. Huebner, R. Dagan, N. Porath, A. D. Wasas, and K. P. Klugman. Lack of utility of

serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. *Pediatr Infect Dis J*, 19(10):1017–20, 2000. URL `https://www.ncbi.nlm.nih.gov/pubmed/11055610`.

[83] K. L. O'Brien, H. Nohynek, and Group World Health Organization Pneumococcal Vaccine Trials Carriage Working. Report from a WHO working group: standard method for detecting upper respiratory carriage of Streptococcus pneumoniae. *Pediatr Infect Dis J*, 22(2):133–40, 2003. doi: 10.1097/01.inf.0000048676.93549.d1. URL `https://www.ncbi.nlm.nih.gov/pubmed/12586977`.

[84] M. Lipsitch. Bacterial vaccines and serotype replacement: lessons from Haemophilus influenzae and prospects for Streptococcus pneumoniae. *Emerg Infect Dis*, 5(3):336–45, 1999. doi: 10.3201/eid0503.990304. URL `https://www.ncbi.nlm.nih.gov/pubmed/10341170`.

[85] N. Lloyd-Evans, T. J. D. O'Dempsey, I. Baldeh, O. Secka, E. Demba, J. E. Todd, T. F. McArdle, W. S. Banya, and B. M. Greenwood. Nasopharyngeal carriage of pneumococci in Gambian children and in their families. *The Pediatric Infectious Disease Journal*, 15(10), 1996. ISSN 0891-3668. URL `https://journals.lww.com/pidj/Fulltext/1996/10000/Nasopharyngeal_carriage_of_pneumococci_in_Gambian.7.aspx`.

[86] M. S. Kaltoft, U. B. Skov Sorensen, H. C. Slotved, and H. B. Konradsen. An easy method for detection of nasopharyngeal carriage of multiple Streptococcus pneumoniae serotypes. *J Microbiol Methods*, 75(3):540–4, 2008. doi: 10.1016/j.mimet.2008.08.010. URL `https://www.ncbi.nlm.nih.gov/pubmed/18801391`.

[87] A. L. Wyllie, Y. Pannekoek, S. Bovenkerk, J. van Engelsdorp Gastelaars, B. Ferwerda, D. van de Beek, E. A. M. Sanders, K. Trzciński, and A. van der Ende. Sequencing of the variable region of rpsB to discriminate between Streptococcus pneumoniae and other streptococcal species. *Open Biol*, 7(9), 2017. ISSN 2046-2441. doi: 10.1098/rsob.170074.

[88] D. A. Tavares, S. Handem, R. J. Carvalho, A. C. Paulo, H. de Lencastre, J. Hinds, and

R. Sá-Leão. Identification of Streptococcus pneumoniae by a real-time PCR assay targeting SP2020. *Scientific Reports*, 9(1):3285, 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-39791-1. URL `https://doi.org/10.1038/s41598-019-39791-1`.

[89] E. Sadowy and W. Hryniewicz. Identification of Streptococcus pneumoniae and other mitis streptococci: importance of molecular methods. *European Journal of Clinical Microbiology I& Infectious Diseases*, 2020. ISSN 1435-4373. doi: 10.1007/s10096-020-03991-9. URL `https://doi.org/10.1007/s10096-020-03991-9`.

[90] S. D. Bentley, D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill, and B. G. Spratt. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet*, 2(3):e31, 2006. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.0020031. URL `https://www.ncbi.nlm.nih.gov/pubmed/16532061`.

[91] A. Fenoll, I. Jado, D. Vicioso, and J. Casal. Dot blot assay for the serotyping of pneumococci. *Journal of clinical microbiology*, 35(3):764–766, 1997. ISSN 0095-1137 1098-660X. doi: 10.1128/JCM.35.3.764-766.1997. URL `https://pubmed.ncbi.nlm.nih.gov/9041430https://www.ncbi.nlm.nih.gov/pmc/articles/PMC229668/`.

[92] M.A. Bronsdon, K.L. O'Brien, R.R. Facklam, C.G. Whitney, B. Schwartz, and G.M. Carlone. Immunoblot method to detect Streptococcus pneumoniae and identify multiple serotypes from nasopharyngeal secretions. *Journal of clinical microbiology*, 42(4): 1596–1600, 2004. doi: 10.1128/jcm.42.4.1596-1600.2004. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC387534/`.

[93] M. C. Enright and B. G. Spratt. A multilocus sequence typing scheme for Streptococcus pneumoniae: identification of clones associated with serious invasive disease. *Microbiology (Reading)*, 144 (Pt 11):3049–60, 1998. doi: 10.1099/00221287-144-11-3049. URL `https://www.ncbi.nlm.nih.gov/pubmed/9846740`.

[94] E. J. Feil, B. C. Li, D. M. Aanensen, W. P. Hanage, and B. G. Spratt. eburst: inferring

patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol*, 186(5):1518–30, 2004. doi: 10.1128/jb.186.5.1518-1530.2004. URL `https://www.ncbi.nlm.nih.gov/pubmed/14973027`.

[95] N. J. Croucher, S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. van der Linden, L. McGee, A. von Gottberg, J. H. Song, K. S. Ko, B. Pichon, S. Baker, C. M. Parry, L. M. Lambertsen, D. Shahinas, D. R. Pillai, T. J. Mitchell, G. Dougan, A. Tomasz, K. P. Klugman, J. Parkhill, W. P. Hanage, and S. D. Bentley. Rapid pneumococcal evolution in response to clinical interventions. *Science*, 331(6016):430–4, 2011. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1198545. URL `https://www.ncbi.nlm.nih.gov/pubmed/21273480`.

[96] L. Epping, A. J. van Tonder, R. A. Gladstone, Consortium The Global Pneumococcal Sequencing, S. D. Bentley, A. J. Page, and J. A. Keane. SeroBA: rapid high-throughput serotyping of Streptococcus pneumoniae from whole genome sequence data. *Microb Genom*, 4(7), 2018. ISSN 2057-5858 (Electronic) 2057-5858 (Linking). doi: 10.1099/mgen.0.000186. URL `https://www.ncbi.nlm.nih.gov/pubmed/29870330`.

[97] G. Kapatai, C. L. Sheppard, A. Al-Shahib, D. J. Litt, A. P. Underwood, T. G. Harrison, and N. K. Fry. Whole genome sequencing of Streptococcus pneumoniae: development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. *PeerJ*, 4:e2477, 2016. doi: 10.7717/peerj.2477. URL `https://www.ncbi.nlm.nih.gov/pubmed/27672516`.

[98] K. L. O'Brien, L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine, T. Cherian, Hib, and Team Pneumococcal Global Burden of Disease Study. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. *Lancet*, 374(9693):893–902, 2009. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(09)61204-6. URL `https://www.ncbi.nlm.nih.gov/pubmed/19748398`.

[99] Global Burden Disease Results Tool, 2017. URL `http://ghdx.healthdata.org/gbd-results-tool`.

[100] Global Burden Disease Results Tool, 2019. URL `http://ghdx.healthdata.org/gbd-results-tool`.

[101] WHO. Pneumonia, 2019. URL `https://www.who.int/news-room/fact-sheets/detail/pneumonia`.

[102] WHO. Pneumococcal disease, 2017. URL `http://www.who.int/immunization/topics/pneumococcal_disease/en/`.

[103] Control US Centers for Disease and Prevention. Active bacterial core surveillance (abcs) report: Emerging infections program network Streptococcus pneumoniae, 2014, September 28, 2015 2014. URL `http://www.cdc.gov/abcs/reports-findings/survreports/spneu14.pdf`.

[104] S. J. Bliss, K. L. O'Brien, E. N. Janoff, M. F. Cotton, P. Musoke, H. Coovadia, and O. S. Levine. The evidence for using conjugate vaccines to protect HIV-infected children against pneumococcal disease. *Lancet Infect Dis*, 8(1):67–80, 2008. doi: 10.1016/S1473-3099(07)70242-6. URL `https://www.ncbi.nlm.nih.gov/pubmed/17974480`.

[105] C. von Mollendorf, A. von Gottberg, S. Tempia, S. Meiring, L. de Gouveia, V. Quan, S. Lengana, T. Avenant, N. du Plessis, B. Eley, H. Finlayson, G. Reubenson, M. Moshe, K. L. O'Brien, K. P. Klugman, C. G. Whitney, C. Cohen, Respiratory Group for Enteric, and Africa Meningeal Disease Surveillance in South. Increased risk for and mortality from invasive pneumococcal disease in HIV-exposed but uninfected infants aged <1 year in South Africa, 2009-2013. *Clin Infect Dis*, 60(9):1346–56, 2015. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1093/cid/civ059. URL `https://www.ncbi.nlm.nih.gov/pubmed/25645212`.

[106] Control US Centers for Disease and Prevention. Pneumococcal disease / facts about

pneumonia, 2017. URL `https://www.cdc.gov/pneumococcal/about/facts.html`.

[107] D. Weycker, E. Richardson, and G. Oster. Childhood vaccination against pneumococcal otitis media and pneumonia: an analysis of benefits and costs. *Am J Manag Care*, 6(10 Suppl):S526–35, 2000. URL `https://www.ncbi.nlm.nih.gov/pubmed/11183904`.

[108] J.R. Casey, D.G. Adlowitz, and M.E. Pichichero. New patterns in the otopathogens causing acute otitis media six to eight years after introduction of pneumococcal conjugate vaccine. *The Pediatric infectious disease journal*, 29(4):304–309, 2010. ISSN 1532-0987 0891-3668. doi: 10.1097/INF.0b013e3181c1bc48. URL `https://pubmed.ncbi.nlm.nih.gov/19935445https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3959886/`.

[109] F. Jover, J. M. Cuadrado, L. Andreu, S. Martinez, R. Canizares, V. O. de la Tabla, C. Martin, P. Roig, and J. Merino. A comparative study of bacteremic and non-bacteremic pneumococcal pneumonia. *Eur J Intern Med*, 19(1):15–21, 2008. ISSN 1879-0828 (Electronic) 0953-6205 (Linking). doi: 10.1016/j.ejim.2007.03.015. URL `https://www.ncbi.nlm.nih.gov/pubmed/18206596`.

[110] M. A. Marcos, M. T. Jimenez de Anta, J. P. de la Bellacasa, J. Gonzalez, E. Martinez, E. Garcia, J. Mensa, A. de Roux, and A. Torres. Rapid urinary antigen test for diagnosis of pneumococcal community-acquired pneumonia in adults. *Eur Respir J*, 21(2):209–14, 2003. doi: 10.1183/09031936.03.00058802. URL `https://www.ncbi.nlm.nih.gov/pubmed/12608431`.

[111] W. P. Hausdorff, G. Siber, and P. R. Paradiso. Geographical differences in invasive pneumococcal disease rates and serotype frequency in young children. *Lancet*, 357 (9260):950–2, 2001. doi: 10.1016/S0140-6736(00)04222-7. URL `https://www.ncbi.nlm.nih.gov/pubmed/11289365`.

[112] F. Ghaffar, T. Barton, J. Lozano, L. S. Muniz, P. Hicks, V. Gan, N. Ahmad, and Jr.

McCracken, G. H. Effect of the 7-valent pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae in the first 2 years of life. *Clin Infect Dis*, 39(7):930–8, 2004. ISSN 1058-4838. doi: 10.1086/423379.

[113] R. A. Adegbola, R. DeAntonio, P. C. Hill, A. Roca, E. Usuf, B. Hoet, and B. M. Greenwood. Carriage of Streptococcus pneumoniae and other respiratory bacterial pathogens in low and lower-middle income countries: a systematic review and meta-analysis. *PLoS One*, 9(8):e103293, 2014. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0103293. URL https://www.ncbi.nlm.nih.gov/pubmed/25084351.

[114] J. O. Hendley, M. A. Sande, P. M. Stewart, and Jr. Gwaltney, J. M. Spread of Streptococcus pneumoniae in families. I. Carriage rates and distribution of types. *J Infect Dis*, 132 (1):55–61, 1975. doi: 10.1093/infdis/132.1.55. URL https://www.ncbi.nlm.nih.gov/pubmed/1097544.

[115] C. J. Chen, Y. C. Huang, L. H. Su, and T. Y. Lin. Nasal carriage of Streptococcus pneumoniae in healthy children and adults in northern Taiwan. *Diagn Microbiol Infect Dis*, 59(3):265–9, 2007. doi: 10.1016/j.diagmicrobio.2007.05.012. URL https://www.ncbi.nlm.nih.gov/pubmed/17900847.

[116] R. G. Pebody, O. Morgan, Y. Choi, R. George, M. Hussain, and N. Andrews. Use of antibiotics and risk factors for carriage of Streptococcus pneumoniae: a longitudinal household study in the United Kingdom. *Epidemiol Infect*, 137(4):555–61, 2009. doi: 10.1017/S0950268808001143. URL https://www.ncbi.nlm.nih.gov/pubmed/18667110.

[117] P. C. Hill, A. Akisanya, K. Sankareh, Y. B. Cheung, M. Saaka, G. Lahai, B. M. Greenwood, and R. A. Adegbola. Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian villagers. *Clin Infect Dis*, 43(6):673–9, 2006. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1086/506941. URL https://www.ncbi.nlm.nih.gov/pubmed/16912937.

[118] G. A. Mackenzie, A. J. Leach, J. R. Carapetis, J. Fisher, and P. S. Morris. Epidemiology of nasopharyngeal carriage of respiratory bacterial pathogens in children and adults: cross-sectional surveys in a population with high rates of pneumococcal disease. *BMC Infect Dis*, 10:304, 2010. ISSN 1471-2334 (Electronic) 1471-2334 (Linking). doi: 10.1186/1471-2334-10-304. URL https://www.ncbi.nlm.nih.gov/pubmed/20969800.

[119] P. Turner, C. Turner, A. Jankhot, N. Helen, S. J. Lee, N. P. Day, N. J. White, F. Nosten, and D. Goldblatt. A longitudinal study of Streptococcus pneumoniae carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. *PLoS One*, 7(5):e38271, 2012. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0038271. URL https://www.ncbi.nlm.nih.gov/pubmed/22693610.

[120] N. Principi, P. Marchisio, G. C. Schito, and S. Mannelli. Risk factors for carriage of respiratory pathogens in the nasopharynx of healthy children. Ascanius Project Collaborative Group. *Pediatr Infect Dis J*, 18(6):517–23, 1999. doi: 10.1097/00006454-199906000-00008. URL https://www.ncbi.nlm.nih.gov/pubmed/10391181.

[121] B. Lopez, M. D. Cima, F. Vazquez, A. Fenoll, J. Gutierrez, C. Fidalgo, M. Caicoya, and F. J. Mendez. Epidemiological study of Streptococcus pneumoniae carriers in healthy primary-school children. *Eur J Clin Microbiol Infect Dis*, 18(11):771–6, 1999. doi: 10.1007/s100960050399. URL https://www.ncbi.nlm.nih.gov/pubmed/10614950.

[122] D. A. Watson, D. M. Musher, J. W. Jacobson, and J. Verhoef. A brief history of the pneumococcus in biomedical research: A panoply of scientific discovery. *Clinical Infectious Diseases*, 17(5):913–924, 1993. ISSN 10584838. URL http://www.jstor.org/stable/4457464.

[123] G. Laible, B. G. Spratt, and R. Hakenbeck. Interspecies recombinational events during the evolution of altered PBP 2x genes in penicillin-resistant clinical isolates of Streptococcus pneumoniae. *Molecular Microbiology*, 5(8):1993–2002, 1991. ISSN 0950-382X. doi: https://doi.org/10.1111/j.1365-2958.1991.

tb00821.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2958.1991.tb00821.x`.

[124] A. Cassini, L. D. Hogberg, D. Plachouras, A. Quattrocchi, A. Hoxha, G. S. Simonsen, M. Colomb-Cotinat, M. E. Kretzschmar, B. Devleesschauwer, M. Cecchini, D. A. Ouakrim, T. C. Oliveira, M. J. Struelens, C. Suetens, D. L. Monnet, and A. M. R. Collaborative Group Burden of. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect Dis*, 19(1):56–66, 2019. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(18)30605-4. URL `https://www.ncbi.nlm.nih.gov/pubmed/30409683`.

[125] C. Zhao, F. Zhang, Y. Chu, Y. Liu, B. Cao, M. Chen, Y. Yu, K. Liao, L. Zhang, Z. Sun, B. Hu, J. Lei, Z. Hu, X. Zhang, and H. Wang. Phenotypic and genotypic characteristic of invasive pneumococcal isolates from both children and adult patients from a multicenter surveillance in China 2005-2011. *PloS one*, 8(12):e82361–e82361, 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0082361. URL `https://pubmed.ncbi.nlm.nih.gov/24349263`.

[126] J. H. Song, R. Dagan, K. P. Klugman, and B. Fritzell. The relationship between pneumococcal serotypes and antibiotic resistance. *Vaccine*, 30(17):2728–2737, 2012. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2012.01.091. URL `http://www.sciencedirect.com/science/article/pii/S0264410X12001648`.

[127] R. Austrian. Pneumococcal vaccine: Development and prospects. *The American Journal of Medicine*, 67(4):547–549, 1979. ISSN 0002-9343. doi: https://doi.org/10.1016/0002-9343(79)90222-5. URL `http://www.sciencedirect.com/science/article/pii/0002934379902225`.

[128] A.E Wright, W. Parry Morgan, L. Colebrook, and R. W. Dodgson. Observations on prophylactic inoculation against pneumococcus infections, and on the results which have been achieved by it. *The Lancet*, 183(4714):1–10, 1914. ISSN 0140-6736. doi: https://

doi.org/10.1016/S0140-6736(01)56370-9. URL `http://www.sciencedirect.com/science/article/pii/S0140673601563709`.

[129] M. Heidelberger and O. T. Avery. The soluble specific substance of pneumococcus. *J Exp Med*, 38(1):73–9, 1923. doi: 10.1084/jem.38.1.73. URL `https://www.ncbi.nlm.nih.gov/pubmed/19868772`.

[130] T. Francis and W. S. Tillett. Cutaneous reactions in pneumonia. the development of antibodies following the intradermal injection of type-specific polysaccharide. *J Exp Med*, 52(4):573–85, 1930. doi: 10.1084/jem.52.4.573. URL `https://www.ncbi.nlm.nih.gov/pubmed/19869789`.

[131] G. M. Ekwurzel, J. S. Simmons, Louis I. Dublin, and Lloyd D. Felton. Studies on immunizing substances in pneumococci: Viii. report on field tests to determine the prophylactic value of a pneumococcus antigen. *Public Health Reports (1896-1970)*, 53(42): 1877–1893, 1938. ISSN 00946214. doi: 10.2307/4582686. URL `http://www.jstor.org/stable/4582686`.

[132] W. G. Smillie, G. H. Warnock, and H. J. White. A study of a Type I pneumococcus epidemic at the state hospital at Worcester, Mass. *Am J Public Health Nations Health*, 28 (3):293–302, 1938. doi: 10.2105/ajph.28.3.293. URL `https://www.ncbi.nlm.nih.gov/pubmed/18014798`.

[133] C. M. Macleod, R. G. Hodges, M. Heidelberger, and W. G. Bernhard. Prevention of pneumococcal pneumonia by immunization with specific capsular polysaccharides. *J Exp Med*, 82(6):445–65, 1945. URL `https://www.ncbi.nlm.nih.gov/pubmed/19871511`.

[134] J. B. Robbins, R. Austrian, C. J. Lee, S. C. Rastogi, G. Schiffman, J. Henrichsen, P. H. Makela, C. V. Broome, R. R. Facklam, R. H. Tiesjema, and et al. Considerations for formulating the second-generation pneumococcal capsular polysaccharide vaccine with emphasis on the cross-reactive types within groups. *J Infect Dis*, 148(6):1136–59,

1983. doi: 10.1093/infdis/148.6.1136. URL https://www.ncbi.nlm.nih.gov/pubmed/6361173.

[135] D. H. Sniadack, B. Schwartz, H. Lipman, J. Bogaerts, J. C. Butler, R. Dagan, G. Echaniz-Aviles, N. Lloyd-Evans, A. Fenoll, N. I. Girgis, and et al. Potential interventions for the prevention of childhood pneumonia: geographic and temporal differences in serotype and serogroup distribution of sterile site pneumococcal isolates from children–implications for vaccine strategies. *Pediatr Infect Dis J*, 14(6):503–10, 1995. URL https://www.ncbi.nlm.nih.gov/pubmed/7667055.

[136] W. P. Hausdorff, J. Bryant, P. R. Paradiso, and G. R. Siber. Which pneumococcal serogroups cause the most invasive disease: Implications for conjugate vaccine formulation and use, part I. *Clin Infect Dis*, 30(1):100–121, 2000. ISSN 1058-4838. doi: 10.1086/313608. URL http://dx.doi.org/10.1086/313608.

[137] W. P. Hausdorff, J. Bryant, C. Kloek, P. R. Paradiso, and G. R. Siber. The contribution of specific pneumococcal serogroups to different disease manifestations: Implications for conjugate vaccine formulation and use, part II. *Clinical Infectious Diseases*, 30(1):122–140, 01 2000. ISSN 1058-4838. doi: 10.1086/313609. URL https://doi.org/10.1086/313609.

[138] S. Moberley, J. Holden, D. P. Tatham, and R. M. Andrews. Vaccines for preventing pneumococcal infection in adults. *The Cochrane database of systematic reviews*, 2013(1): CD000422–CD000422, 2013. ISSN 1469-493X. doi: 10.1002/14651858.CD000422.pub3. URL https://pubmed.ncbi.nlm.nih.gov/23440780.

[139] G. T. Rijkers, E. A. M. Sanders, M. A. Breukels, and B. J. M. Zegers. Responsiveness of infants to capsular polysaccharides: implications for vaccine development. *Reviews in Medical Microbiology*, 7(1), 1996. ISSN 0954-139X. URL https://journals.lww.com/revmedmicrobiol/Fulltext/1996/01000/Responsiveness_of_infants_to_capsular.1.aspx.

[140] K. L. O'Brien, L. H. Moulton, R. Reid, R. Weatherholtz, J. Oski, L. Brown, G. Kumar,

A. Parkinson, D. Hu, J. Hackell, I. Chang, R. Kohberger, G. Siber, and M. Santosham. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *Lancet*, 362(9381):355–61, 2003. doi: 10. 1016/S0140-6736(03)14022-6. URL `https://www.ncbi.nlm.nih.gov/pubmed/12907008`.

[141] Control US Centers for Disease and Prevention. Progress toward elimination of Haemophilus influenzae type b disease among infants and children–United States, 1987-1993. *MMWR Morb Mortal Wkly Rep*, 43(8):144–8, 1994. URL `https://www.ncbi.nlm.nih.gov/pubmed/8309463`.

[142] K. A. Robinson, W. Baughman, G. Rothrock, N. L. Barrett, M. Pass, C. Lexau, B. Damaske, K. Stefonek, B. Barnes, J. Patterson, E. R. Zell, A. Schuchat, C. G. Whitney, and Network for the Active Bacterial Core Surveillance/Emerging Infections Program. Epidemiology of invasive infections in the United States, 1995-1998: Opportunities for prevention in the conjugate vaccine era. *JAMA*, 285(13): 1729–1735, 2001. ISSN 0098-7484. doi: 10.1001/jama.285.13.1729. URL `https://doi.org/10.1001/jama.285.13.1729`.

[143] H. R. Shinefield, S. Black, P. Ray, I. Chang, N. Lewis, B. Fireman, J. Hackell, P. R. Paradiso, G. Siber, R. Kohberger, D. V. Madore, F. J. Malinowski, A. Kimura, C. Le, I. Landaw, J. Aguilar, and J. Hansen. Safety and immunogenicity of heptavalent pneumococcal CRM197 conjugate vaccine in infants and toddlers. *Pediatr Infect Dis J*, 18(9):757–63, 1999.

[144] M. B. Rennels, K. M. Edwards, H. L. Keyserling, K. S. Reisinger, D. A. Hogerman, D. V. Madore, I. Chang, P. R. Paradiso, F. J. Malinoski, and A. Kimura. Safety and immunogenicity of heptavalent pneumococcal vaccine conjugated to CRM197 in United States infants. *Pediatrics*, 101(4 Pt 1):604–11, 1998. URL `https://www.ncbi.nlm.nih.gov/pubmed/9521941`.

[145] S. Black, H. Shinefield, B. Fireman, E. Lewis, P. Ray, J. R. Hansen, L. Elvin, K. M.

Ensor, J. Hackell, G. Siber, F. Malinoski, D. Madore, I. Chang, R. Kohberger, W. Watson, R. Austrian, and K. Edwards. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. *Pediatr Infect Dis J*, 19(3):187–95, 2000.

[146] C. G. Whitney, M. M. Farley, J. Hadler, L. H. Harrison, C. Lexau, A. Reingold, L. Lefkowitz, P. R. Cieslak, M. Cetron, E. R. Zell, J. H. Jorgensen, and A. Schuchat. Increasing prevalence of multidrug-resistant Streptococcus pneumoniae in the United States. *New England Journal of Medicine*, 343(26):1917–1924, 2000. doi: 10.1056/nejm200012283432603. URL https://www.nejm.org/doi/full/10.1056/NEJM200012283432603.

[147] K. P. Klugman, S. A. Madhi, R. E. Huebner, R. Kohberger, N. Mbelle, N. Pierce, and Group Vaccine Trialists. A trial of a 9-valent pneumococcal conjugate vaccine in children with and those without HIV infection. *N Engl J Med*, 349(14):1341–8, 2003. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/NEJMoa035060. URL https://www.ncbi.nlm.nih.gov/pubmed/14523142.

[148] WHO Expert Committee on Biological Standardization. Recommendations for the production and control of pneumococcal conjugate vaccines. Technical Report February 1, 2021, World Health Organisation, 2005. URL https://www.who.int/biologicals/WHO_TRS_%20927%20full%20tex.pdf.

[149] R. Dagan, R. Melamed, M. Muallem, L. Piglansky, D. Greenberg, O. Abramson, P. M. Mendelman, N. Bohidar, and P. Yagupsky. Reduction of nasopharyngeal carriage of pneumococci during the second year of life by a heptavalent conjugate pneumococcal vaccine. *J Infect Dis*, 174(6):1271–8, 1996. URL https://www.ncbi.nlm.nih.gov/pubmed/8940218.

[150] S. K. Obaro, R. A. Adegbola, W. A. Banya, and B. M. Greenwood. Carriage of pneumococci after pneumococcal vaccination. *Lancet*, 348(9022):271–2, 1996. URL https://www.ncbi.nlm.nih.gov/pubmed/8684225.

[151] N. Mbelle, R. E. Huebner, A. D. Wasas, A. Kimura, I. Chang, and K. P. Klugman. Immunogenicity and impact on nasopharyngeal carriage of a nonavalent pneumococcal conjugate vaccine. *J Infect Dis*, 180(4):1171–6, 1999. doi: 10.1086/315009. URL `https://www.ncbi.nlm.nih.gov/pubmed/10479145`.

[152] C. G. Whitney, M. M. Farley, J. Hadler, L. H. Harrison, N. M. Bennett, R. Lynfield, A. Reingold, P. R. Cieslak, T. Pilishvili, D. Jackson, R. R. Facklam, J. H. Jorgensen, A. Schuchat, and Network Active Bacterial Core Surveillance of the Emerging Infections Program. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med*, 348(18):1737–46, 2003. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/NEJMoa022823. URL `https://www.ncbi.nlm.nih.gov/pubmed/12724479`.

[153] D. F. Vestrheim, Ø. Løvoll, I. S. Aaberge, D. A. Caugant, E. A. Høiby, H. Bakke, and M. R. Bergsaker. Effectiveness of a 2+1 dose schedule pneumococcal conjugate vaccination programme on invasive pneumococcal disease among children in Norway. *Vaccine*, 26(26):3277–3281, 2008. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2008.03.087. URL `http://www.sciencedirect.com/science/article/pii/S0264410X08004210`.

[154] M. van der Linden, G. Falkenhorst, S. Perniciaro, C. Fitzner, and M. Imöhl. Effectiveness of pneumococcal conjugate vaccines (PCV7 and PCV13) against invasive pneumococcal disease among children under two years of age in Germany. *PLOS ONE*, 11(8):e0161257, 2016. doi: 10.1371/journal.pone.0161257. URL `https://doi.org/10.1371/journal.pone.0161257`.

[155] J. Benito-Fernández, S. M. Raso, I. Pocheville-Gurutzeta, J. SánchezEtxaniz, B. Azcunaga-Santibañez, and S. Capapé-Zache. Pneumococcal bacteremia among infants with fever without known source before and after introduction of pneumococcal conjugate vaccine in the Basque Country of Spain. *The Pediatric Infectious Disease Journal*, 26(8), 2007. ISSN 0891-

3668. URL `https://journals.lww.com/pidj/Fulltext/2007/08000/Pneumococcal_Bacteremia_Among_Infants_With_Fever.2.aspx`.

[156] F. Dubos, I. Marechal, M. O. Husson, C. Courouble, M. Aurel, and A. Martinot. Decline in pneumococcal meningitis after the introduction of the heptavalent-pneumococcal conjugate vaccine in northern France. *Archives of Disease in Childhood*, 92(11):1009–1012, 2007. doi: 10.1136/adc.2007.119891. URL `https://adc.bmj.com/content/archdischild/92/11/1009.full.pdf`.

[157] X. Li, C. Mukandavire, Z. M. Cucunubá, S. Echeverria Londono, K. Abbas, H. E. Clapham, M. Jit, H. L. Johnson, T. Papadopoulos, E. Vynnycky, M. Brisson, E. D. Carter, A. Clark, M. J. de Villiers, K. Eilertson, M. J. Ferrari, I. Gamkrelidze, K. A. M. Gaythorpe, N. C. Grassly, T. B. Hallett, W. Hinsley, M. L. Jackson, K. Jean, A. Karachaliou, P. Klepac, J. Lessler, X. Li, S. M. Moore, S. Nayagam, D. M. Nguyen, H. Razavi, D. Razavi-Shearer, S. Resch, C. Sanderson, S. Sweet, S. Sy, Y. Tam, H. Tanvir, Q. M. Tran, C. L. Trotter, S. Truelove, K. van Zandvoort, S. Verguet, N. Walker, A. Winter, K. Woodruff, N. M. Ferguson, and T. Garske. Estimating the health impact of vaccination against ten pathogens in 98 low-income and middle-income countries from 2000 to 2030: a modelling study. *The Lancet*, 397(10272):398–408, 2021. ISSN 0140-6736. doi: 10.1016/S0140-6736(20)32657-X. URL `https://doi.org/10.1016/S0140-6736(20)32657-X`.

[158] WHO. Pneumococcal vaccines WHOposition paper–2012. *Wkly Epidemiol Rec*, 87(14):129–44, 2012. URL `https://www.ncbi.nlm.nih.gov/pubmed/24340399`.

[159] L. Conklin, J. D. Loo, J. Kirk, K. E. Fleming-Dutra, M. Deloria Knoll, D. E. Park, D. Goldblatt, K. L. O'Brien, and C. G. Whitney. Systematic review of the effect of pneumococcal conjugate vaccine dosing schedules on vaccine-type invasive pneumococcal disease among young children. *Pediatr Infect Dis J*, 33 Suppl 2:S109–18, 2014. ISSN 1532-0987 (Electronic) 0891-3668 (Linking). doi: 10.1097/INF.0000000000000078. URL `https://www.ncbi.nlm.nih.gov/pubmed/24336053`.

[160] M. R. Griffin, Y. Zhu, M. R. Moore, C. G. Whitney, and C. G. Grijalva. U.S. hospital-

izations for pneumonia after a decade of pneumococcal vaccination. *N Engl J Med*, 369(2):155–63, 2013. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/ NEJMoa1209165. URL `https://www.ncbi.nlm.nih.gov/pubmed/23841730`.

[161] C. Levy, E. Varon, C. Picard, S. Bechet, A. Martinot, S. Bonacorsi, and R. Cohen. Trends of pneumococcal meningitis in children after introduction of the 13-valent pneumococcal conjugate vaccine in France. *Pediatr Infect Dis J*, 33(12):1216–21, 2014. ISSN 1532-0987 (Electronic) 0891-3668 (Linking). doi: 10.1097/INF.0000000000000451. URL `https://www.ncbi.nlm.nih.gov/pubmed/25037044`.

[162] S. Ben-Shimol, D. Greenberg, N. Givon-Lavi, Y. Schlesinger, E. Somekh, S. Aviner, D. Miron, and R. Dagan. Early impact of sequential introduction of 7-valent and 13-valent pneumococcal conjugate vaccine on ipd in Israeli children <5 years: an active prospective nationwide surveillance. *Vaccine*, 32(27):3452–9, 2014. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2014.03.065. URL `https://www.ncbi.nlm.nih.gov/pubmed/24690148`.

[163] Z. B. Harboe, T. Dalby, D. M. Weinberger, T. Benfield, K. Molbak, H. C. Slotved, C. H. Suppli, H. B. Konradsen, and P. Valentiner-Branth. Impact of 13-valent pneumococcal conjugate vaccination in invasive pneumococcal disease incidence and mortality. *Clin Infect Dis*, 59(8):1066–73, 2014. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10. 1093/cid/ciu524. URL `https://www.ncbi.nlm.nih.gov/pubmed/25034421`.

[164] N. J. Andrews, P. A. Waight, P. Burbidge, E. Pearce, L. Roalfe, M. Zancolli, M. Slack, S. N. Ladhani, E. Miller, and D. Goldblatt. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *Lancet Infect Dis*, 14(9):839–46, 2014. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(14)70822-9. URL `https://www.ncbi.nlm.nih. gov/pubmed/25042756`.

[165] C. L. F. Walker, I. Rudan, L. Liu, H. Nair, E. Theodoratou, Z. A. Bhutta, K. L. O'Brien, H. Campbell, and R. E. Black. Global burden of childhood pneumonia and diarrhoea.

Lancet, 381(9875):1405–1416, 2013. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(13)60222-6. URL https://www.ncbi.nlm.nih.gov/pubmed/23582727.

[166] S. M. Davis, M. Deloria-Knoll, H. T. Kassa, and K. L. O'Brien. Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: review of evidence on indirect effects. *Vaccine*, 32(1):133–45, 2013. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2013.05.005. URL https://www.ncbi.nlm.nih.gov/pubmed/23684824.

[167] E. Miller, N. J. Andrews, P. A. Waight, M. P. Slack, and R. C. George. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. *Lancet Infect Dis*, 11(10):760–8, 2011. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(11)70090-1. URL https://www.ncbi.nlm.nih.gov/pubmed/21621466.

[168] WHO. Changing epidemiology of pneumococcal serotypes after introduction of conjugate vaccine: July 2010 report. *Wkly Epidemiol Rec*, 85(43):434–6, 2010. URL https://www.ncbi.nlm.nih.gov/pubmed/21038716.

[169] W. P. Hanage, K. Auranen, R. Syrjanen, E. Herva, P. H. Makela, T. Kilpi, and B. G. Spratt. Ability of pneumococcal serotypes and clones to cause acute otitis media: implications for the prevention of otitis media by conjugate vaccines. *Infect Immun*, 72(1):76–81, 2004.

[170] B. G. Spratt and B. M. Greenwood. Prevention of pneumococcal disease by vaccination: does serotype replacement matter? *The Lancet*, 356(9237):1210–1211, 2000. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(00)02779-3. URL http://www.sciencedirect.com/science/article/pii/S0140673600027793.

[171] D. F. Vestrheim, E. A. Høiby, I. S. Aaberge, and D. A. Caugant. Impact of a pneumococcal conjugate vaccination program on carriage among children in Norway. *Clinical and*

vaccine immunology : CVI, 17(3):325–334, 2010. doi: 10.1128/CVI.00435-09. URL `https://pubmed.ncbi.nlm.nih.gov/20107006`.

[172] R. Sa-Leao, S. Nunes, A. Brito-Avo, N. Frazao, A. S. Simoes, M. I. Crisostomo, A. C. Paulo, J. Saldanha, I. Santos-Sanches, and H. de Lencastre. Changes in pneumococcal serotypes and antibiotypes carried by vaccinated and unvaccinated day-care centre attendees in Portugal, a country with widespread use of the seven-valent pneumococcal conjugate vaccine. *Clin Microbiol Infect*, 15(11):1002–7, 2009. ISSN 1469-0691 (Electronic) 1198-743X (Linking). doi: 10.1111/j.1469-0691.2009.02775.x. URL `https://www.ncbi.nlm.nih.gov/pubmed/19392883`.

[173] D. M. Weinberger, R. Malley, and M. Lipsitch. Serotype replacement in disease after pneumococcal vaccination. *Lancet*, 378(9807):1962–73, 2011. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(10)62225-8. URL `https://www.ncbi.nlm.nih.gov/pubmed/21492929`.

[174] P. C. Wroe, G. M. Lee, J. A. Finkelstein, S. I. Pelton, W. P. Hanage, M. Lipsitch, A. E. Stevenson, S. L. Rifas-Shiman, K. Kleinman, M. M. Dutta-Linn, V. L. Hinrichsen, M. Lakoma, and S. S. Huang. Pneumococcal carriage and antibiotic resistance in young children before 13-valent conjugate vaccine. *The Pediatric infectious disease journal*, 31(3):249–254, 2012. doi: 10.1097/INF.0b013e31824214ac. URL `https://pubmed.ncbi.nlm.nih.gov/22173142https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3288953/`.

[175] M. Lipsitch. Interpreting results from trials of pneumococcal conjugate vaccines: A statistical test for detecting vaccine-induced increases in carriage of nonvaccine serotypes. *American Journal of Epidemiology*, 154(1):85–92, 2001. ISSN 0002-9262. doi: 10.1093/aje/154.1.85. URL `https://doi.org/10.1093/aje/154.1.85`.

[176] N. Frazão, R. Sá-Leão, and H. de Lencastre. Impact of a single dose of the 7-valent pneumococcal conjugate vaccine on colonization. *Vaccine*, 28(19):3445–3452, 2010.

ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2010.02.070. URL `http://www.sciencedirect.com/science/article/pii/S0264410X10002513`.

[177] S. D. Brugger, P. Frey, S. Aebi, J. Hinds, and K. Muhlemann. Multiple colonization with S. pneumoniae before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine. *PLoS One*, 5(7):e11638, 2010. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0011638. URL `https://www.ncbi.nlm.nih.gov/pubmed/20661289`.

[178] S. S. Joshi, M. A. Al-Mamun, and D. M. Weinberger. Correlates of nonrandom patterns of serotype switching in pneumococcus. *J Infect Dis*, 221(10):1669–1676, 2020. ISSN 1537-6613 (Electronic) 0022-1899 (Linking). doi: 10.1093/infdis/jiz687. URL `https://www.ncbi.nlm.nih.gov/pubmed/31875229`.

[179] N. J. Croucher, L. Kagedan, C. M. Thompson, J. Parkhill, S. D. Bentley, J. A. Finkelstein, M. Lipsitch, and W. P. Hanage. Selective and genetic constraints on pneumococcal serotype switching. *PLoS Genet*, 11(3):e1005095, 2015. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.1005095. URL `https://www.ncbi.nlm.nih.gov/pubmed/25826208`.

[180] K. L. Wyres, L. M. Lambertsen, N. J. Croucher, L. McGee, A. von Gottberg, J. Linares, M. R. Jacobs, K. G. Kristinsson, B. W. Beall, K. P. Klugman, J. Parkhill, R. Hakenbeck, S. D. Bentley, and A. B. Brueggemann. Pneumococcal capsular switching: a historical perspective. *J Infect Dis*, 207(3):439–49, 2013. ISSN 1537-6613 (Electronic) 0022-1899 (Linking). doi: 10.1093/infdis/jis703. URL `https://www.ncbi.nlm.nih.gov/pubmed/23175765`.

[181] W. P. Hanage, C. J. Bishop, S. S. Huang, A. E. Stevenson, S. I. Pelton, M. Lipsitch, and J. A. Finkelstein. Carried pneumococci in Massachusetts children: the contribution of clonal expansion and serotype switching. *Pediatr Infect Dis J*, 30(4):302–8, 2011. ISSN 1532-0987 (Electronic) 0891-3668 (Linking). doi: 10.1097/INF.0b013e318201a154. URL `https://www.ncbi.nlm.nih.gov/pubmed/21085049`.

[182] N. J. Croucher, C. Chewapreecha, W. P. Hanage, S. R. Harris, L. McGee, M. van der Linden, J. H. Song, K. S. Ko, H. de Lencastre, C. Turner, F. Yang, R. Sa-Leao, B. Beall, K. P. Klugman, J. Parkhill, P. Turner, and S. D. Bentley. Evidence for soft selective sweeps in the evolution of pneumococcal multidrug resistance and vaccine escape. *Genome Biol Evol*, 6(7):1589–602, 2014. ISSN 1759-6653 (Electronic) 1759-6653 (Linking). doi: 10.1093/gbe/evu120. URL https://www.ncbi.nlm.nih.gov/pubmed/24916661.

[183] N. J. Croucher, J. A. Finkelstein, S. I. Pelton, P. K. Mitchell, G. M. Lee, J. Parkhill, S. D. Bentley, W. P. Hanage, and M. Lipsitch. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nature Genetics*, 45(6):656–663, 2013. ISSN 1546-1718. doi: 10.1038/ng.2625. URL https://doi.org/10.1038/ng.2625.

[184] J. Corander, C. Fraser, M. U. Gutmann, B. Arnold, W. P. Hanage, S. D. Bentley, M. Lipsitch, and N. J. Croucher. Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nature Ecology & Evolution*, 1(12):1950–1960, 2017. ISSN 2397-334X. doi: 10.1038/s41559-017-0337-x. URL https://doi.org/10.1038/s41559-017-0337-x.

[185] E. Leibovitz, M. R. Jacobs, and R. Dagan. Haemophilus influenzae: a significant pathogen in acute otitis media. *Pediatr Infect Dis J*, 23(12):1142–52, 2004.

[186] M. W. Tregnaghi, X. Saez-Llorens, P. Lopez, H. Abate, E. Smith, A. Posleman, A. Calvo, D. Wong, C. Cortes-Barbosa, A. Ceballos, M. Tregnaghi, A. Sierra, M. Rodriguez, M. Troitino, C. Carabajal, A. Falaschi, A. Leandro, M. M. Castrejon, A. Lepetic, P. Lommel, W. P. Hausdorff, D. Borys, J. Ruiz Guinazu, E. Ortega-Barria, J. P. Yarzabal, L. Schuerman, and Compas Group. Efficacy of pneumococcal nontypable Haemophilus influenzae protein D conjugate vaccine (PHiD-CV) in young Latin American children: A double-blind randomized controlled trial. *PLoS Med*, 11(6):e1001657, 2014. ISSN 1549-1676 (Electronic) 1549-1277 (Linking). doi: 10.1371/journal.pmed.1001657. URL https://www.ncbi.nlm.nih.gov/pubmed/24892763.

[187] R. Prymula, P. Peeters, V. Chrobok, P. Kriz, E. Novakova, E. Kaliskova, I. Kohl, P. Lom-

mel, J. Poolman, J. P. Prieels, and L. Schuerman. Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both Streptococcus pneumoniae and non-typable Haemophilus influenzae: a randomised double-blind efficacy study. *Lancet*, 367(9512):740–8, 2006. ISSN 0140-6736. doi: 10.1016/s0140-6736(06)68304-9.

[188] S. Ruckinger, R. von Kries, A. Siedler, and M. van der Linden. Association of serotype of Streptococcus pneumoniae with risk of severe and fatal outcome. *Pediatr Infect Dis J*, 28(2):118–22, 2009. doi: 10.1097/INF.0b013e318187e215. URL `https://www.ncbi.nlm.nih.gov/pubmed/19116604`.

[189] G. L. Rodgers and K. P. Klugman. The future of pneumococcal disease prevention. *Vaccine*, 29 Suppl 3:C43–8, 2011. ISSN 0264-410x. doi: 10.1016/j.vaccine.2011.07.047.

[190] J. Linares, C. Ardanuy, R. Pallares, and A. Fenoll. Changes in antimicrobial resistance, serotypes and genotypes in Streptococcus pneumoniae over a 30-year period. *Clin Microbiol Infect*, 16(5):402–10, 2010. ISSN 1469-0691 (Electronic) 1198-743X (Linking). doi: 10.1111/j.1469-0691.2010.03182.x. URL `https://www.ncbi.nlm.nih.gov/pubmed/20132251`.

[191] K. Ampofo, A.T. Pavia, C. Stockmann, A.L. Hersh, J.M. Bender, A.J. Blaschke, H.Y.C. Weng, K.E. Korgenski, J. Daly, E.O. Mason, and C.L. Byington. The changing epidemiology of invasive pneumococcal disease at a tertiary children's hospital through the PCV7 era. *The Pediatric infectious disease journal*, 31(3):10.1097/INF.0b013e31823dcc72, 2012. ISSN 0891-3668 1532-0987. doi: 10.1097/INF.0b013e31823dcc72. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3299810/`.

[192] D. Sharma, W. Baughman, A. Holst, S. Thomas, D. Jackson, M. da Gloria Carvalho, B. Beall, S. Satola, R. Jerris, S. Jain, M. M. Farley, and J. P. Nuorti. Pneumococcal carriage and invasive disease in children before introduction of the 13-valent conjugate vaccine: comparison with the era before 7-valent conjugate vaccine. *Pediatr Infect Dis J*, 32(2): e45–53, 2013. ISSN 0891-3668. doi: 10.1097/INF.0b013e3182788fdd.

[193] C. M. Giele, A. D. Keil, D. Lehmann, and P. G. Van Buynder. Invasive pneumococcal disease in Western Australia: emergence of serotype 19A. *Med J Aust*, 190(3):166, 2009. doi: 10.5694/j.1326-5377.2009.tb02332.x. URL `https://www.ncbi.nlm.nih.gov/pubmed/19203324`.

[194] J. A. Bettinger, D. W. Scheifele, J. D. Kellner, S. A. Halperin, W. Vaudry, B. Law, G. Tyrrell, and Active Canadian Immunization Monitoring Program. The effect of routine vaccination on invasive pneumococcal infections in canadian children, immunization monitoring program, active 2000-2007. *Vaccine*, 28(9):2130–6, 2010. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2009.12.026. URL `https://www.ncbi.nlm.nih.gov/pubmed/20044050`.

[195] G. Hanquet, E. Kissling, A. Fenoll, R. George, A. Lepoutre, T. Lernout, D. Tarrago, E. Varon, and J. Verhaegen. Pneumococcal serotypes in children in 4 European countries. *Emerg Infect Dis*, 16(9):1428–39, 2010. ISSN 1080-6059 (Electronic) 1080-6040 (Linking). doi: 10.3201/eid1609.100102. URL `https://www.ncbi.nlm.nih.gov/pubmed/20735928`.

[196] D. Ozawa, H. Yano, S. Endo, H. Hidaka, R. Kakuta, N. Okitsu, H. Kanamori, M. Ogawa, S. Ichimura, M. Shimojima, S. Inomata, M. Kaku, and Y. Katori. Impact of the seven-valent pneumococcal conjugate vaccine on acute otitis media in Japanese children: Emergence of serotype 15a multidrug-resistant Streptococcus pneumoniae in middle ear fluid isolates. *Pediatr Infect Dis J*, 34(9):e217–21, 2015. doi: 10.1097/INF.0000000000000776. URL `https://www.ncbi.nlm.nih.gov/pubmed/26083590`.

[197] M. C. Brandileone, S. C. G. Almeida, R. Minamisava, and A. L. Andrade. Distribution of invasive Streptococcus pneumoniae serotypes before and 5years after the introduction of 10-valent pneumococcal conjugate vaccine in Brazil. *Vaccine*, 36(19):2559–2566, 2018. doi: 10.1016/j.vaccine.2018.04.010. URL `https://www.ncbi.nlm.nih.gov/pubmed/29650385`.

[198] T. Pilishvili, C. Lexau, M. M. Farley, J. Hadler, L. H. Harrison, N. M. Bennett, A. Rein-

gold, A. Thomas, W. Schaffner, A. S. Craig, P. J. Smith, B. W. Beall, C. G. Whitney, M. R. Moore, and Network Active Bacterial Core Surveillance/Emerging Infections Program. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis*, 201(1):32–41, 2010. ISSN 1537-6613 (Electronic) 0022-1899 (Linking). doi: 10.1086/648593. URL https://www.ncbi.nlm.nih.gov/pubmed/19947881.

[199] R. Dagan, N. Givon-Lavi, O. Zamir, and D. Fraser. Effect of a nonavalent conjugate vaccine on carriage of antibiotic-resistant Streptococcus pneumoniae in day-care centers. *Pediatr Infect Dis J*, 22(6):532–40, 2003. doi: 10.1097/01.inf.0000069761.11093.c3. URL https://www.ncbi.nlm.nih.gov/pubmed/12799510.

[200] D. S. Stephens, S. M. Zughaier, C. G. Whitney, W. S. Baughman, L. Barker, K. Gay, D. Jackson, W. A. Orenstein, K. Arnold, A. Schuchat, M. M. Farley, and Program Georgia Emerging Infections. Incidence of macrolide resistance in Streptococcus pneumoniae after introduction of the pneumococcal conjugate vaccine: population-based assessment. *Lancet*, 365(9462):855–63, 2005. ISSN 1474-547X (Electronic) 0140-6736 (Linking). doi: 10.1016/S0140-6736(05)71043-6. URL https://www.ncbi.nlm.nih.gov/pubmed/15752529.

[201] B. Temple, N. T. Toan, V. T. T. Dai, K. Bright, P. V. Licciardi, R. A. Marimla, C. D. Nguyen, D. Y. Uyen, A. Balloch, T. N. Huu, and E. Kim Mulholland. Immunogenicity and reactogenicity of ten-valent versus 13-valent pneumococcal conjugate vaccines among infants in Ho Chi Minh City, Vietnam: a randomised controlled trial. *The Lancet Infectious Diseases*, 19(5):497–509, 2019. ISSN 1473-3099. doi: https://doi.org/10.1016/S1473-3099(18)30734-5. URL http://www.sciencedirect.com/science/article/pii/S1473309918307345.

[202] A. J. Wijmenga-Monsuur, E. van Westen, M. J. Knol, R. M. C. Jongerius, M. Zancolli, D. Goldblatt, P. G. M. van Gageldonk, I. Tcherniaeva, G. A. M. Berbers, and N. Y. Rots. Direct comparison of immunogenicity induced by 10- or 13-valent pneumococcal conjugate vaccine around the 11-month booster in Dutch infants. *PloS One*, 10(12):e0144739–0144739, 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.

0144739. URL `https://pubmed.ncbi.nlm.nih.gov/26658902https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4690595/`.

[203] R. Lapidot, K. M. Shea, I. Yildirim, H. J. Cabral, S. I. Pelton, and The Massachusetts Department Of Public Health. Characteristics of serotype 3 invasive pneumococcal disease before and after universal childhood immunization with PCV13 in Massachusetts. *Pathogens (Basel, Switzerland)*, 9(5):396, 2020. ISSN 2076-0817. doi: 10.3390/pathogens9050396. URL `https://pubmed.ncbi.nlm.nih.gov/32455770https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7281000/`.

[204] M. Kremer, J. D. Levin, and C. M. Snyder. Advance market commitments: Insights from theory and experience. Working Paper 26775, National Bureau of Economic Research, February 2020. URL `http://www.nber.org/papers/w26775`.

[205] IVAC. VIEW-hub, January 11, 2021 2021. URL `https://view-hub.org`.

[206] E. Clarke, A. O. Bashorun, M. Okoye, A. Umesi, M. Badjie Hydara, I. Adigweme, R. Dhere, V. Sethna, B. Kampmann, D. Goldblatt, A. Tate, D. H. Weiner, J. Flores, M. R. Alderson, and S. Lamola. Safety and immunogenicity of a novel 10-valent pneumococcal conjugate vaccine candidate in adults, toddlers, and infants in The Gambia-results of a phase 1/2 randomized, double-blinded, controlled trial. *Vaccine*, 38(2):399–410, 2020. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2019.08.072. URL `https://www.ncbi.nlm.nih.gov/pubmed/31843266`.

[207] M. E. Pichichero. Pneumococcal whole-cell and protein-based vaccines: changing the paradigm. *Expert review of vaccines*, 16(12):1181–1190, 2017. ISSN 1744-8395 1476-0584. doi: 10.1080/14760584.2017.1393335. URL `https://pubmed.ncbi.nlm.nih.gov/29130395https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6277969/`.

[208] C. Gianfaldoni, S. Censini, M. Hilleringmann, M. Moschioni, C. Facciotti, W. Pansegrau, V. Masignani, A. Covacci, R. Rappuoli, M. A. Barocchi, and P. Ruggiero. Streptococcus pneumoniae pilus subunits protect mice against lethal challenge. *Infection and Immunity*,

75(2):1059–1062, 2007. doi: 10.1128/iai.01400-06. URL `https://iai.asm.org/content/iai/75/2/1059.full.pdf`.

[209] G. Regev-Yochay, W. P. Hanage, K. Trzcinski, S. L. Rifas-Shiman, G. Lee, A. Bessolo, S. S. Huang, S. I. Pelton, A. J. McAdam, J. A. Finkelstein, M. Lipsitch, and R. Malley. Re-emergence of the type 1 pilus among Streptococcus pneumoniae isolates in Massachusetts, USA. *Vaccine*, 28(30):4842–6, 2010. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2010.04.042. URL `https://www.ncbi.nlm.nih.gov/pubmed/20434550`.

[210] A. Basset, K.H. Turner, E. Boush, S. Sayeed, S.L. Dove, and R. Malley. Expression of the type 1 pneumococcal pilus is bistable and negatively regulated by the structural component rrga. *Infection and Immunity*, 79(8):2974–2983, 2011. doi: 10.1128/iai.05117-11. URL `https://iai.asm.org/content/iai/79/8/2974.full.pdf`.

[211] R.M. Anderson and R.M. May. *Infectious diseases of humans: dynamics and control*. Oxford Science Publications. Oxford University Press, Oxford, 1991. ISBN 0-19-854599-1.

[212] M. Jit and M. Brisson. Modelling the epidemiology of infectious diseases for decision analysis: a primer. *Pharmacoeconomics*, 29(5):371–86, 2011. ISSN 1179-2027 (Electronic) 1170-7690 (Linking). doi: 10.2165/11539960-000000000-00000. URL `https://www.ncbi.nlm.nih.gov/pubmed/21504239`.

[213] N. C. Grassly and C. Fraser. Mathematical models of infectious disease transmission. *Nat Rev Microbiol*, 6(6):477–87, 2008. ISSN 1740-1534 (Electronic) 1740-1526 (Linking). doi: 10.1038/nrmicro1845. URL `https://www.ncbi.nlm.nih.gov/pubmed/18533288`.

[214] M. Roberts, V. Andreasen, A. Lloyd, and L. Pellis. Nine challenges for deterministic epidemic models. *Epidemics*, 10:49–53, 2015. ISSN 1878-0067 (Electronic) 1878-0067 (Linking). doi: 10.1016/j.epidem.2014.09.006. URL `https://www.ncbi.nlm.nih.gov/pubmed/25843383`.

[215] L. J. Allen and A. M. Burgin. Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Math Biosci*, 163(1):1–33, 2000. doi: 10.1016/s0025-5564(99) 00047-4. URL `https://www.ncbi.nlm.nih.gov/pubmed/10652843`.

[216] J. A. Jacquez and C. P. Simon. The stochastic SI model with recruitment and deaths. I. Comparison with the closed SIS model. *Math Biosci*, 117(1-2):77–125, 1993. doi: 10.1016/0025-5564(93)90018-6. URL `https://www.ncbi.nlm.nih.gov/pubmed/8400585`.

[217] J. P. Fox, L. Elveback, W. Scott, L. Gatewood, and E. Ackerman. Herd immunity: basic concept and relevance to public health immunization practices. *Am J Epidemiol*, 94 (3):179–89, 1971. doi: 10.1093/oxfordjournals.aje.a121310. URL `https://www.ncbi.nlm.nih.gov/pubmed/5093648`.

[218] R.M. Anderson. *Chapter 1 - The Impact of Vaccination on the Epidemiology of Infectious Diseases*, pages 3–31. Academic Press, 2016. ISBN 978-0-12-802174-3. doi: https://doi.org/10.1016/B978-0-12-802174-3.00001-1. URL `https://www.sciencedirect.com/science/article/pii/B9780128021743000011`.

[219] V. E. Pitzer and N. E. Basta. Linking data and models the importance of statistical analyses to inform models for the transmission dynamics of infections. *Epidemiology*, 23(4):520–522, 2012. ISSN 1044-3983. doi: 10.1097/EDE.0b013e31825902ab.

[220] E. Gjini. Geographic variation in pneumococcal vaccine efficacy estimated from dynamic modeling of epidemiological data post-PCV7. *Sci Rep*, 7(1):3049, 2017. ISSN 2045-2322 (Electronic) 2045-2322 (Linking). doi: 10.1038/s41598-017-02955-y. URL `https://www.ncbi.nlm.nih.gov/pubmed/28607461`.

[221] E. Gjini and M. G. M. Gomes. Expanding vaccine efficacy estimation with dynamic models fitted to cross-sectional prevalence data post-licensure. *Epidemics*, 14:71–82, 2016. ISSN 1755-4365. doi: http://dx.doi.org/10.1016/j.epidem.2015.11.001. URL `http://www.sciencedirect.com/science/article/pii/S1755436515000948`.

[222] S. Flasche, O. Le Polain de Waroux, K. L. O'Brien, and W. J. Edmunds. The serotype distribution among healthy carriers before vaccination is essential for predicting the impact of pneumococcal conjugate vaccine on invasive disease. *PLoS Comput Biol*, 11 (4):e1004173, 2015. ISSN 1553-734x. doi: 10.1371/journal.pcbi.1004173.

[223] D. M. Weinberger, L. R. Grant, R. C. Weatherholtz, J. L. Warren, K. L. O'Brien, and L. L. Hammitt. Relating pneumococcal carriage among children to disease rates among adults before and after the introduction of conjugate vaccines. *Am J Epidemiol*, 183(11): 1055–62, 2016. ISSN 0002-9262. doi: 10.1093/aje/kwv283.

[224] S. E. Talbird, A. S. Ismaila, and T. N. Taylor. A steady-state, population-based model to estimate the direct and indirect effects of pneumococcal vaccines. *Vaccine*, 28 Suppl 6: G3–13, 2010. ISSN 0264-410x. doi: 10.1016/j.vaccine.2010.06.013.

[225] B. Standaert, N. Demarteau, S. Talbird, and J. Mauskopf. Modelling the effect of conjugate vaccines in pneumococcal disease: cohort or population models? *Vaccine*, 28 Suppl 6:G30–8, 2010. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine. 2010.06.015. URL https://www.ncbi.nlm.nih.gov/pubmed/21075268.

[226] P. De Wals, S. Black, R. Borrow, and D. Pearce. Modeling the impact of a new vaccine on pneumococcal and nontypable Haemophilus influenzae diseases: a new simulation model. *Clin Ther*, 31(10):2152–69, 2009. ISSN 0149-2918. doi: 10.1016/j.clinthera.2009. 10.014.

[227] F. J. Zhou, M. H. Kyaw, A. Shefer, C. A. Winston, and J. P. Nuorti. Health care utilization for pneumonia in young children after routine pneumococcal conjugate vaccine use in the United States. *Archives of Pediatrics & Adolescent Medicine*, 161(12):1162–1168, 2007. ISSN 1072-4710. doi: 10.1001/archpedi.161.12.1162.

[228] M. Paulden. Recent amendments to nice's value-based assessment of health technologies: implicitly inequitable? *Expert Review of Pharmacoeconomics and Outcomes Research*, 17(3):239–242, 2017. ISSN 1473-7167. doi: 10.1080/14737167.2017.1330152. URL https://doi.org/10.1080/14737167.2017.1330152.

[229] M.Y. Bertram, J.A. Lauer, K. De Joncheere, T. Edejer, R. Hutubessy, MP. Kieny, and S.R. Hill. Cost-effectiveness thresholds: pros and cons. *Bulletin of the World Health Organization*, 94(12):925–930, 2016. ISSN 1564-0604 0042-9686. doi: 10.2471/BLT.15. 164418. URL `https://pubmed.ncbi.nlm.nih.gov/27994285`.

[230] M. Drummond. *Methods for the economic evaluation of health care programmes*. Oxford medical publications. Oxford University Press, Oxford ; New York, 3rd edition, 2005. ISBN 0198529449. URL `http://www.loc.gov/catdir/enhancements/fy0640/2005279427-d.html`.

[231] A. Brennan, S. E. Chick, and R. Davies. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ*, 15(12):1295–310, 2006. doi: 10.1002/hec. 1148. URL `https://www.ncbi.nlm.nih.gov/pubmed/16941543`.

[232] J. Caro, D.M. Eddy, H. Kan, C. Kaltz, B. Patel, R. Eldessouki, and A.H. Briggs. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: An ISPOR-AMCP-NPC good practice task force report. *Value in Health*, 17(2):174–182, 2014. ISSN 1098-3015. doi: https://doi.org/10.1016/j.jval. 2014.01.003. URL `http://www.sciencedirect.com/science/article/pii/S1098301514000114`.

[233] B. F. M. Wijnen, Gapg Van Mastrigt, W. K. Redekop, H. J. M. Majoie, R. J. A. De Kinderen, and Smaa Evers. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: data extraction, risk of bias, and transferability (part 3/3). *Expert Review of Pharmacoeconomics and Outcomes Research*, 16(6):723–732, 2016. ISSN 1473-7167. doi: 10.1080/14737167.2016.1246961. URL `https://doi.org/10.1080/14737167.2016.1246961`.

[234] C. Bottomley, A. Roca, P. C. Hill, B. Greenwood, and V. Isham. A mathematical model of serotype replacement in pneumococcal carriage following vaccination. *J R Soc Interface*, 10(89):20130786, 2013. ISSN 1742-5662. doi: 10.1098/rsif.2013.0786.

[235] Y. H. Choi, M. Jit, N. Gay, N. Andrews, P. A. Waight, A. Melegaro, R. George, and

E. Miller. 7-valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? *PLoS One*, 6(10):e26190, 2011. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0026190. URL https://www.ncbi.nlm.nih.gov/pubmed/22022559.

[236] Y. H. Choi, N. Andrews, and E. Miller. Estimated impact of revising the 13-valent pneumococcal conjugate vaccine schedule from 2+1 to 1+1 in England and Wales: A modelling study. *PLoS Med*, 16(7):e1002845–e1002845, 2019. ISSN 1549-1676 1549-1277. doi: 10.1371/journal.pmed.1002845. URL https://www.ncbi.nlm.nih.gov/pubmed/31269018https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6608946/.

[237] E. De Cao, A. Melegaro, R. Klok, and M. Postma. Optimising assessments of the epidemiological impact in the Netherlands of paediatric immunisation with 13-valent pneumococcal conjugate vaccine using dynamic transmission modelling. *Plos One*, 9 (4), 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0089415.

[238] S. Flasche, J. Ojal, O. Le Polain de Waroux, M. Otiende, K. L. O'Brien, M. Kiti, D. J. Nokes, W. J. Edmunds, and J. A. G. Scott. Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Medicine*, 15(1): 113, 2017. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=prem&AN=28592303.

[239] E. Gjini, C. Valente, R. Sá-Leão, and M. G. M. Gomes. How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *Journal of Theoretical Biology*, 388:50–60, 2016. ISSN 0022-5193. doi: http://dx.doi.org/10.1016/j.jtbi.2015.09.031. URL http://www.sciencedirect.com/science/article/pii/S0022519315004816.

[240] T. Kitano and H. Aoki. The incremental burden of invasive pneumococcal disease associated with a decline in childhood vaccination using a dynamic transmission model in Japan: A secondary impact of covid-19. *Comput Biol Med*, 133:104429, 2021. ISSN

1879-0534 (Electronic) 0010-4825 (Linking). doi: 10.1016/j.compbiomed.2021.104429. URL https://www.ncbi.nlm.nih.gov/pubmed/33930765.

[241] O. Le Polain De Waroux, W. J. Edmunds, K. Takahashi, K. Ariyoshi, E. K. Mulholland, D. Goldblatt, Y. H. Choi, D. D. Anh, L. M. Yoshida, and S. Flasche. Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccin Immunother*, 14(8):1939–1947, 2018. ISSN 2164-5515. doi: 10.1080/21645515.2018. 1467201.

[242] M. Lipsitch. Vaccination against colonizing bacteria with multiple serotypes. *Proc Natl Acad Sci U S A*, 94(12):6571–6, 1997. URL https://www.ncbi.nlm.nih.gov/pubmed/9177259.

[243] A. Melegaro, Y. H. Choi, R. George, W. J. Edmunds, E. Miller, and N. J. Gay. Dynamic models of pneumococcal carriage and the impact of the heptavalent pneumococcal conjugate vaccine on invasive pneumococcal disease. *Bmc Infectious Diseases*, 10, 2010. ISSN 1471-2334. doi: 10.1186/1471-2334-10-90.

[244] J. Ojal, S. Flasche, L. L. Hammitt, D. Akech, M. C. Kiti, T. Kamau, I. Adetifa, M. Nurhonen, J. A. G. Scott, and K. Auranen. Sustained reduction in vaccine-type invasive pneumococcal disease despite waning effects of a catch-up campaign in Kilifi, Kenya: A mathematical model based on pre-vaccination data. *Vaccine*, 2017. ISSN 0264-410x. doi: 10.1016/j.vaccine.2017.07.019.

[245] J. Ojal, U. Griffiths, L. L. Hammitt, I. Adetifa, D. Akech, C. Tabu, J. A. G. Scott, and S. Flasche. Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *The Lancet Global Health*, 7(5):e644–e654, 2019. ISSN 2214-109X. doi: https://doi.org/10.1016/S2214-109X(18)30562-X. URL http://www.sciencedirect.com/science/article/pii/S2214109X1830562X.

[246] R. Omori, B. J. Cowling, and H. Nishiura. How is vaccine effectiveness scaled by the

transmission dynamics of interacting pathogen strains with cross-protective immunity? *Plos One*, 7(11), 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0050751.

[247] P. Scott, S. A. Herzog, K. Auranen, R. Dagan, N. Low, M. Egger, and J. C. M. Heijne. Timing of bacterial carriage sampling in vaccine trials: A modelling study. *Epidemics*, 9:8–17, 2014. ISSN 1755-4365. doi: http://dx.doi.org/10.1016/j.epidem.2014.08.003. URL `http://www.sciencedirect.com/science/article/pii/S1755436514000383`.

[248] S. J. Snedecor, D. R. Strutton, V. Ciuryla, E. J. Schwartz, and M. F. Botteman. Transmission-dynamic model to capture the indirect effects of infant vaccination with Prevnar (7-valent pneumococcal conjugate vaccine (PCV7)) in older populations. *Vaccine*, 27(34):4694–4703, 2009. ISSN 0264-410X. doi: 10.1016/j.vaccine.2009.05.062. URL `<GotoISI>://WOS:000268351200019`.

[249] L. Temime, D. Guillemot, and P. Y. Boelle. Short- and long-term effects of pneumococcal conjugate vaccination of children on penicillin resistance. *Antimicrob Agents Chemother*, 48(6):2206–13, 2004. doi: 10.1128/aac.48.6.2206-2213.2004.

[250] T. Van Effelterre, M. R. Moore, F. Fierens, C. G. Whitney, L. White, S. I. Pelton, and W. P. Hausdorff. A dynamic model of pneumococcal infection in the United States: Implications for prevention through vaccination. *Vaccine*, 28(21):3650–3660, 2010. ISSN 0264-410X. doi: 10.1016/j.vaccine.2010.03.030. URL `<GotoISI>://WOS:000278128800006`.

[251] M. Wasserman, A. Lucas, D. Jones, M. Wilson, B. Hilton, A. Vyse, H. Madhava, A. Brogan, M. Slack, and R. Farkouh. Dynamic transmission modelling to address infant pneumococcal conjugate vaccine schedule modifications in the UK. *Epidemiol Infect*, 146(14):1797–1806, 2018. ISSN 0950-2688. doi: 10.1017/s095026881800198x.

[252] Y. Zhang, K. Auranen, and M. Eichner. The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. *Epidemiology and Infection*, 132

(6):1073–1081, 2004. ISSN 0950-2688. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2870198/`.

[253] P. K. Mitchell, M. Lipsitch, and W. P. Hanage. Carriage burden, multiple colonization and antibiotic pressure promote emergence of resistant vaccine escape pneumococci. *Philos Trans R Soc Lond B Biol Sci*, 370(1670):20140342, 2015. ISSN 0962-8436. doi: 10.1098/rstb.2014.0342.

[254] S. Flasche, W. J. Edmunds, E. Miller, D. Goldblatt, C. Robertson, and Y. H. Choi. The impact of specific and non-specific immunity on the ecology of Streptococcus pneumoniae and the implications for vaccination. *Proc Biol Sci*, 280(1771):20131939, 2013. ISSN 1471-2954 (Electronic) 0962-8452 (Linking). doi: 10.1098/rspb.2013.1939. URL `https://www.ncbi.nlm.nih.gov/pubmed/24089337`.

[255] M. Nurhonen, A. C. Cheng, and K. Auranen. Pneumococcal transmission and disease in silico: a microsimulation model of the indirect effects of vaccination. *PLoS One*, 8(2): e56079, 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0056079.

[256] A. Yang, F. Cai, and M. Lipsitch. Herd immunity alters the conditions for performing dose schedule comparisons: an individual-based model of pneumococcal carriage. *BMC Infect Dis*, 19(1):227, 2019. ISSN 1471-2334. doi: 10.1186/s12879-019-3833-6.

[257] Y. H. Choi, M. Jit, S. Flasche, N. Gay, and E. Miller. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. *Plos One*, 7(7), 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0039927. URL `<GotoISI>://WOS:000306406700015`.

[258] F. Ansaldi, S. Pugh, D. Amicizia, R. Di Virgilio, C. Trucchi, A. Orsi, A. Zollo, and G. Icardi. Estimating the clinical and economic impact of switching from the 13-valent pneumococcal conjugate vaccine (PCV13) to the 10-valent pneumococcal conjugate vaccine (PCV10) in Italy. *Pathogens*, 9(2), 2020. doi: 10.3390/pathogens9020076.

[259] M. Bakir, O. Turel, and O. Topachevskyi. Cost-effectiveness of new pneumococcal conjugate vaccines in Turkey: a decision analytical model. *BMC health services research*, 12:386,

2012. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed14&NEWS=N&AN=366406373`.

[260] P. R. Blank and T. D. Szucs. Cost-effectiveness of 13-valent pneumococcal conjugate vaccine in switzerland. *Vaccine*, 30(28):4267–4275, 2012. ISSN 0264-410X. doi: http://dx.doi.org/10.1016/j.vaccine.2012.04.028. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed14&NEWS=N&AN=51992016`.

[261] A. W. Chuck, P. Jacobs, G. Tyrrell, and J. D. Kellner. Pharmacoeconomic evaluation of 10- and 13-valent pneumococcal conjugate vaccines. *Vaccine*, 28(33):5485–90, 2010. ISSN 0264-410x. doi: 10.1016/j.vaccine.2010.05.058.

[262] P. De Wals, B. Poirier, G. Petit, L. Erickson, and J. Pepin. Simulation model for comparing the costs and effectiveness of different pneumococcal conjugate vaccines. *Procedia in Vaccinology*, 1:67–72, 2009. doi: 10.1016/j.provac.2009.07.012. URL `http://www.sciencedirect.com/science/article/pii/S1877282X09000435`.

[263] S. R. Earnshaw, C. L. McDade, G. Zanotti, R. A. Farkouh, and D. Strutton. Cost-effectiveness of 2+1 dosing of 13-valent and 10-valent pneumococcal conjugate vaccines in Canada. *BMC Infectious Diseases*, 12(101), 2012. ISSN 1471-2334. URL `http://www.biomedcentral.com/content/pdf/1471-2334-12-101.pdf`.

[264] R. M. Klok, R. M. Lindkvist, M. Ekelund, R. A. Farkouh, and D. R. Strutton. Cost-effectiveness of a 10- versus 13-valent pneumococcal conjugate vaccine in Denmark and Sweden. *Clinical Therapeutics*, 35(2):119–134, 2013. ISSN 0149-2918. doi: http://dx.doi.org/10.1016/j.clinthera.2012.12.006. URL `http://www.sciencedirect.com/science/article/pii/S014929181200673X`.

[265] M. M. Nakamura, A. Tasslimi, T. A. Lieu, O. Levine, M. D. Knoll, L. B. Russell, and A. Sinha. Cost effectiveness of child pneumococcal conjugate vaccination in middle-income countries. *International Health (RSTMH)*, 3(4):270–281, 2011. ISSN 1876-3413. doi: http://dx.doi.org/10.1016/j.inhe.2011.08.004. URL `https://academic.oup.com/inthealth/article-lookup/doi/10.1016/j.inhe.2011.08.004`.

[266] S. Pugh, M. Wasserman, M. Moffatt, S. Marques, J. M. Reyes, V. A. Prieto, D. Reijnders, M. H. Rozenbaum, J. Laine, H. Åhman, and R. Farkouh. Estimating the impact of switching from a lower to higher valent pneumococcal conjugate vaccine in Colombia, Finland, and The Netherlands: A cost-effectiveness analysis. *Infectious Diseases and Therapy*, 9(2):305–324, 2020. doi: 10.1007/s40121-020-00287-5. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7237584/`.

[267] D. R. Strutton, R. A. Farkouh, S. R. Earnshaw, S. Hwang, U. Theidel, S. Kontodimas, R. Klok, and S. Papanicolaou. Cost-effectiveness of 13-valent pneumococcal conjugate vaccine: Germany, Greece, and The Netherlands. *Journal of Infection*, 64(1):54–67, 2012. ISSN 0163-4453. doi: 10.1016/j.jinf.2011.10.015. URL `<GotoISI>://WOS:000298635400006`.

[268] S. E. Talbird, T. N. Taylor, S. Knoll, C. R. Frostad, and S. G. Marti. Outcomes and costs associated with PHiD-CV, a new protein D conjugate pneumococcal vaccine, in four countries. *Vaccine*, 28(SUPPL. 6):G23–G29, 2010. ISSN 0264-410X. doi: http://dx.doi.org/10.1016/j.vaccine.2010.06.016. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed12&NEWS=N&AN=359956839`.

[269] M. H. Rozenbaum, E. A. M. Sanders, A. J. Van Hoek, A. G. S. C. Jansen, A. Van Der Ende, G. Van Den Dobbelsteen, G. D. Rodenburg, E. Hak, and M. J. Postma. Cost effectiveness of pneumococcal vaccination among Dutch infants: An economic analysis of the seven valent pneumococcal conjugated vaccine and forecast for the 10 valent and 13 valent vaccines. *BMJ (Online)*, 340(7760):1347, 2010. doi: http://dx.doi.org/10.1136/bmj.c2509. URL `http://www.bmj.com/cgi/reprint/340/jun02_1/c2509http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed12&NEWS=N&AN=359028551`.

[270] V. V. Vucina, S. K. Filipovic, N. Koznjak, V. Stamenic, A. D. Clark, B. Mounaud, J. Blau, C. Hoestlandt, and B. Kaic. Cost-effectiveness of pneumococcal conjugate vaccination in Croatia. *Vaccine*, 33:A209–A218, 2015. ISSN 0264-410X. doi: 10.1016/j.vaccine.2014.12.043. URL `<GotoISI>://WOS:000355034500032`.

[271] P. Vemer and M. J. Postma. A few years later: update of the cost-effectiveness of infant pneumococcal vaccination in Dutch children. *Human Vaccines and Immunotherapeutics*, 10(7):1841–1849, 2014. ISSN 2164-5515. doi: 10.4161/hv.29008. URL `<GotoISI>: //WOS:000342901200018`.

[272] M. Wilson, M. Wasserman, T. Jadavi, M. Postma, M. C. Breton, F. Peloquin, S. Earnshaw, C. McDade, H. Sings, and R. Farkouh. Clinical and economic impact of a potential switch from 13-valent to 10-valent pneumococcal conjugate infant vaccination in Canada. *Infectious Diseases and Therapy*, 2018. ISSN 2193-6382. doi: 10.1007/s40121-018-0206-1. URL `https://doi.org/10.1007/s40121-018-0206-1`.

[273] A. By, P. Sobocki, A. Forsgren, and S. A. Silfverdal. Comparing health outcomes and costs of general vaccination with pneumococcal conjugate vaccines in Sweden: a Markov model. *Clinical Therapeutics*, 34(1):177–189, 2012. ISSN 0149-2918. doi: http://dx.doi.org/10.1016/j.clinthera.2011.12.007. URL `http://www.sciencedirect.com/science/article/pii/S0149291811008423`.

[274] P. Castiglia, L. Pradelli, S. Castagna, V. Freguglia, G. Palu, and S. Esposito. Overall effectiveness of pneumococcal conjugate vaccines: an economic analysis of PHiD-CV and PCV13 in the immunization of infants in Italy. *Hum Vaccin Immunother*, page 0, 2017. ISSN 2164-5515. doi: 10.1080/21645515.2017.1343773.

[275] E. Delgleize, O. Leeuwenkamp, E. Theodorou, and N. van de Velde. Cost-effectiveness analysis of routine pneumococcal vaccination in the UK: a comparison of the PHiD-CV vaccine and the PCV-13 vaccine using a Markov model. *BMJ Open*, 6(11):e010776, 2016. ISSN 2044-6055. doi: http://dx.doi.org/10.1136/bmjopen-2015-010776. URL `http://bmjopen.bmj.com/content/6/11/e010776`.

[276] G. Knerer, A. Ismaila, and D. Pearce. Health and economic impact of PHiD-CV in Canada and the UK: A Markov modelling exercise. *Journal of Medical Economics*, 15(1):61–76, 2012. ISSN 1369-6998. doi: http://dx.doi.org/10.3111/

13696998.2011.622323. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS&` `PAGE=reference&D=emed14&NEWS=N&AN=364117371`.

[277] B. Robberstad, C. R. Frostad, P. E. Akselsen, K. J. Kvaerner, and A. K. H. Berstad. Economic evaluation of second generation pneumococcal conjugate vaccines in Norway. *Vaccine*, 29(47):8564–8574, 2011. ISSN 0264-410X. doi: 10.1016/j.vaccine.2011.09.025.

[278] P. Beutels, A. Blommaert, G. Hanquet, J. Bilcke, N. Thiry, M. Sabbe, J. Verhaegen, F. De Smet, M. Callens, and P. Van Damme. Cost-effectiveness of 10- and 13-valent pneumococcal conjugate vaccines in childhood. *KCE Reports*, 155C, 2011. URL `https://kce.fgov.be/sites/default/files/page_documents/kce_` `155c_pneumococcal_vaccines.pdf`.

[279] K. J. Smith, A. R. Wateska, M. P. Nowalk, L. H. Harrison, R. K. Zimmerman, C. J. Lin, and W. Schaffner. Higher-valency pneumococcal conjugate vaccines: An exploratory cost-effectiveness analysis in U.S. seniors. *American Journal of Preventive Medicine*, 61 (1):28–36, 2021. ISSN 0749-37971873-2607. doi: http://dx.doi.org/10.1016/j.amepre. 2021.01.023. URL `https://www.sciencedirect.com/science/article/pii/` `S0749379721001239#tb1fn1`.

[280] A. Kuhlmann and J. M. G. von der Schulenburg. Modeling the cost-effectiveness of infant vaccination with pneumococcal conjugate vaccines in Germany. *European Journal of Health Economics*, 18(3):273–292, 2017. ISSN 1618-7598. doi: http://dx.doi.org/10.1007/s10198-016-0770-9. URL `http://link.springer.de/` `link/service/journals/10198/index.htmhttp://ovidsp.ovid.com/` `ovidweb.cgi?T=JS&PAGE=reference&D=emex&NEWS=N&AN=608645357`.

[281] J. L. Rubin, L. J. McGarry, D. R. Strutton, K. P. Klugman, S. I. Pelton, K. E. Gilmore, and M. C. Weinstein. Public health and economic impact of the 13-valent pneumococcal conjugate vaccine (PCV13) in the United States. *Vaccine*, 28(48):7634–7643, 2010. ISSN 0264-410X. doi: 10.1016/j.vaccine.2010.09.049.

[282] D. L. Luca, J. C. Kwong, A. Chu, B. Sander, R. O'Reilly, A. J. McGeer, and D. E. Bloom.

Impact of pneumococcal vaccination on pneumonia hospitalizations and related costs in Ontario: A population-based ecological study. *Clin Infect Dis*, 66(4):541–547, 2018. doi: 10.1093/cid/cix850.

[283] A. Waye and A. W. Chuck. Value added by the Prevnar 13 childhood immunization program in Alberta, Canada (2010-2015). *Drugs - Real World Outcomes*, 2(3):311–318, 2015. ISSN 2199-1154. doi: http://dx.doi.org/10.1007/s40801-015-0037-2. URL `http://www.springer.com/adis/journal/40801`.

[284] International Vaccine Access Center (IVAC). View-hub report: Global vaccine introduction and implementation, 2017.

[285] GlaxoSmithKline. Synflorix™, GlaxoSmithKline's pneumococcal vaccine, receives European authorisation, 2009. URL `http://www.gsk.com/en-gb/media/press-releases/synflorix-glaxosmithkline-s-pneumococcal-vaccine-receives-european-auth`

[286] FDA. Prevnar 13, 2017. URL `https://www.fda.gov/biologicsbloodvaccines/vaccines/approvedproducts/ucm201667.htm`.

[287] S. J. Snedecor, M. F. Botteman, and D. R. Strutton. Serotype-specific transmission dynamics of invasive pneumococcal disease after vaccination with 7-valent pneumococcal conjugate vaccine. *Value in Health*, 12 (7):A393, 2009. URL `http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed12&AN=70002829`.

[288] K. Auranen, J. Mehtala, A. Tanskanen, and S. Kaltoft M. Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. *Am J Epidemiol*, 171(2):169–76, 2010. ISSN 1476-6256 (Electronic) 0002-9262 (Linking). doi: 10.1093/aje/kwp351. URL `https://www.ncbi.nlm.nih.gov/pubmed/19969530`.

[289] D. Pessoa, F. Hoti, R. Syrjanen, R. Sa-Leao, T. Kaijalainen, M. G. Gomes, and K. Auranen. Comparative analysis of Streptococcus pneumoniae transmission in Portuguese and

Finnish day-care centres. *BMC Infect Dis*, 13:180, 2013. ISSN 1471-2334 (Electronic) 1471-2334 (Linking). doi: 10.1186/1471-2334-13-180. URL `https://www.ncbi.nlm.nih.gov/pubmed/23597389`.

[290] M. Lipsitch, C. Colijn, T. Cohen, W. P. Hanage, and C. Fraser. No coexistence for free: neutral null models for multistrain pathogens. *Epidemics*, 1(1):2–13, 2009. ISSN 1878-0067 (Electronic) 1878-0067 (Linking). doi: 10.1016/j.epidem.2008.07.001. URL `https://www.ncbi.nlm.nih.gov/pubmed/21352747`.

[291] E. Numminen, L. Cheng, M. Gyllenberg, and J. Corander. Estimating the transmission dynamics of Streptococcus pneumoniae from strain prevalence data. *Biometrics*, 69(3): 748–57, 2013. ISSN 1541-0420 (Electronic) 0006-341X (Linking). doi: 10.1111/biom.12040. URL `https://www.ncbi.nlm.nih.gov/pubmed/23822205`.

[292] M. Krol and W. Brouwer. Unpaid work in health economic evaluations. *Soc Sci Med*, 144:127–37, 2015. ISSN 0277-9536. doi: 10.1016/j.socscimed.2015.09.008.

[293] L. S. Nymark, T. Sharma, A. Miller, U. Enemark, and U. K. Griffiths. Inclusion of the value of herd immunity in economic evaluations of vaccines. a systematic review of methods used. *Vaccine*, 35(49, Part B):6828–6841, 2017. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2017.10.024. URL `http://www.sciencedirect.com/science/article/pii/S0264410X17314202`.

[294] M. K. Siggins, S. K. Gill, P. R. Langford, Y. Li, S. N. Ladhani, and J. S. Tregoning. PHiD-CV induces anti-protein D antibodies but does not augment pulmonary clearance of nontypeable Haemophilus influenzae in mice. *Vaccine*, 33(38):4954–61, 2015. ISSN 0264-410x. doi: 10.1016/j.vaccine.2015.07.034.

[295] M. R. van den Bergh, J. Spijkerman, K. M. Swinnen, N. A. Francois, T. G. Pascal, D. Borys, L. Schuerman, E. P. Ijzerman, J. P. Bruin, A. van der Ende, R. H. Veenhoven, and E. A. Sanders. Effects of the 10-valent pneumococcal nontypeable Haemophilus influenzae protein D-conjugate vaccine on nasopharyngeal bacterial colonization in

young children: a randomized controlled trial. *Clin Infect Dis*, 56(3):e30–9, 2013. ISSN 1058-4838. doi: 10.1093/cid/cis922.

[296] R. Dagan, S.I. Pelton, L. Bakaletz, and R. Cohen. Prevention of early episodes of otitis media by pneumococcal vaccines might reduce progression to complex disease. *The Lancet Infectious Diseases*, 16(4):480–492, 2016. ISSN 1473-3099. doi: https://doi.org/10.1016/S1473-3099(15)00549-6. URL http://www.sciencedirect.com/science/article/pii/S1473309915005496.

[297] C. G. Whitney, T. Pilishvili, M. M. Farley, W. Schaffner, A. S. Craig, R. Lynfield, A. C. Nyquist, K. A. Gershman, M. Vazquez, N. M. Bennett, A. Reingold, A. Thomas, M. P. Glode, E. R. Zell, J. H. Jorgensen, B. Beall, and A. Schuchat. Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. *The Lancet*, 368(9546):1495–1502, 2006. ISSN 0140-6736. doi: http://dx.doi.org/10.1016/S0140-6736(06)69637-2. URL http://www.sciencedirect.com/science/article/pii/S0140673606696372.

[298] C. M. Domingues, J. R. Verani, E. I. Montenegro Renoiner, M. C. de Cunto Brandileone, B. Flannery, L. H. de Oliveira, J. B. Santos, and J. C. de Moraes. Effectiveness of ten-valent pneumococcal conjugate vaccine against invasive pneumococcal disease in Brazil: a matched case-control study. *Lancet Respir Med*, 2(6):464–71, 2014. ISSN 2213-2600. doi: 10.1016/s2213-2600(14)70060-8.

[299] Belgian Directorate-General Healthcare, 2015. URL http://www.health.belgium.be/internet2Prd/groups/public/@public/@shc/documents/ie2ddiver/19103438.pdf.

[300] G. Deceuninck, G. De Serres, N. Boulianne, B. Lefebvre, and P. De Wals. Effectiveness of three pneumococcal conjugate vaccines to prevent invasive pneumococcal disease in Quebec, Canada. *Vaccine*, 33(23):2684–9, 2015. ISSN 0264-410x. doi: 10.1016/j.vaccine.2015.04.005.

[301] J. Jokinen, H. Rinta-Kokko, L. Siira, A. A. Palmu, M. J. Virtanen, H. Nohynek,

A. Virolainen-Julkunen, M. Toropainen, and J. P. Nuorti. Impact of ten-valent pneumo-coccal conjugate vaccination on invasive pneumococcal disease in finnish children – a population-based study. *PLOS ONE*, 10(3):e0120290, 2015. doi: 10.1371/journal.pone. 0120290. URL `https://doi.org/10.1371/journal.pone.0120290`.

[302] Z. B. Harboe, P. Valentiner-Branth, T. L. Benfield, J. J. Christensen, P. H. Andersen, M. Howitz, K. A. Krogfelt, L. Lambertsen, and H. B. Konradsen. Early effectiveness of heptavalent conjugate pneumococcal vaccination on invasive pneumococcal disease af-ter the introduction in the danish childhood immunization programme. *Vaccine*, 28(14): 2642–7, 2010. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine. 2010.01.017. URL `https://www.ncbi.nlm.nih.gov/pubmed/20096392`.

[303] F. Dorléans, E. Varon, A. Lepoutre, S. Georges, L. Gutmann, and D. Lévy-Bruhl, 2010. URL `http://invs.santepubliquefrance.fr/presse/2010/le_point_sur/vaccination_pneumo_050710/index.html`.

[304] G. Deceuninck, P. De Wals, N. Boulianne, and G. De Serres. Effectiveness of pneu-mococcal conjugate vaccine using a 2+1 infant schedule in Quebec, Canada. *Pediatr Infect Dis J*, 29(6):546–9, 2010. ISSN 1532-0987 (Electronic) 0891-3668 (Linking). doi: 10.1097/INF.0b013e3181cffa2a. URL `https://www.ncbi.nlm.nih.gov/pubmed/20125062`.

[305] E. D. McIntosh and R. R. Reinert. Global prevailing and emerging pediatric pneumo-coccal serotypes. *Expert Rev Vaccines*, 10(1):109–29, 2011. ISSN 1744-8395 (Electronic) 1476-0584 (Linking). doi: 10.1586/erv.10.145. URL `https://www.ncbi.nlm.nih.gov/pubmed/21162625`.

[306] D. M. Weinberger, K. Trzcinski, Y. J. Lu, D. Bogaert, A. Brandes, J. Galagan, P. W. Anderson, R. Malley, and M. Lipsitch. Pneumococcal capsular polysaccharide struc-ture predicts serotype prevalence. *PLoS Pathog*, 5(6):e1000476, 2009. ISSN 1553-7374 (Electronic) 1553-7366 (Linking). doi: 10.1371/journal.ppat.1000476. URL `https://www.ncbi.nlm.nih.gov/pubmed/19521509`.

[307] C. C. Sanders, Jr. Sanders, W. E., and D. J. Harrowe. Bacterial interference: effects of oral antibiotics on the normal throat flora and its ability to interfere with group a streptococci. *Infect Immun*, 13(3):808–12, 1976. URL https://www.ncbi.nlm.nih.gov/pubmed/818017.

[308] Jr. Johanson, W. G., R. Blackstock, A. K. Pierce, and J. P. Sanford. The role of bacterial antagonism in pneumococcal colonization of the human pharynx. *J Lab Clin Med*, 75(6): 946–52, 1970. URL https://www.ncbi.nlm.nih.gov/pubmed/4393022.

[309] S. Dawid, A. M. Roche, and J. N. Weiser. The blp bacteriocins of Streptococcus pneumoniae mediate intraspecies competition both in vitro and in vivo. *Infect Immun*, 75(1): 443–51, 2007. doi: 10.1128/IAI.01775-05. URL https://www.ncbi.nlm.nih.gov/pubmed/17074857.

[310] S. Guiral, T. J. Mitchell, B. Martin, and J. P. Claverys. Competence-programmed predation of noncompetent cells in the human pathogen Streptococcus pneumoniae: genetic requirements. *Proc Natl Acad Sci U S A*, 102(24):8710–5, 2005. doi: 10.1073/pnas.0500879102. URL https://www.ncbi.nlm.nih.gov/pubmed/15928084.

[311] L. J. Hathaway, S. D. Brugger, B. Morand, M. Bangert, J. U. Rotzetter, C. Hauser, W. A. Graber, S. Gore, A. Kadioglu, and K. Muhlemann. Capsule type of Streptococcus pneumoniae determines growth phenotype. *PLoS Pathog*, 8(3):e1002574, 2012. ISSN 1553-7374 (Electronic) 1553-7366 (Linking). doi: 10.1371/journal.ppat.1002574. URL https://www.ncbi.nlm.nih.gov/pubmed/22412375.

[312] A. L. Nelson, A. M. Roche, J. M. Gould, K. Chim, A. J. Ratner, and J. N. Weiser. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun*, 75(1):83–90, 2007. doi: 10.1128/IAI.01475-06. URL https://www.ncbi.nlm.nih.gov/pubmed/17088346.

[313] M. Carrolo, M. J. Frias, F. R. Pinto, J. Melo-Cristino, and M. Ramirez. Prophage spontaneous activation promotes dna release enhancing biofilm formation in Streptococcus pneumoniae. *PLoS One*, 5(12):e15678, 2010. ISSN 1932-6203 (Electronic) 1932-6203

(Linking). doi: 10.1371/journal.pone.0015678. URL `https://www.ncbi.nlm.nih.gov/pubmed/21187931`.

[314] M. Lipsitch, J. K. Dykes, S. E. Johnson, E. W. Ades, J. King, D. E. Briles, and G. M. Carlone. Competition among Streptococcus pneumoniae for intranasal colonization in a mouse model. *Vaccine*, 18(25):2895–901, 2000. URL `https://www.ncbi.nlm.nih.gov/pubmed/10812233`.

[315] M. Brisson and W. J. Edmunds. Economic evaluation of vaccination programs: the impact of herd-immunity. *Med Decis Making*, 23(1):76–82, 2003. doi: 10.1177/0272989x02239651.

[316] H. C. Turner, M. Walker, M. D. French, I. M. Blake, T. S. Churcher, and M. G. Basanez. Neglected tools for neglected diseases: mathematical models in economic evaluations. *Trends Parasitol*, 30(12):562–70, 2014. ISSN 1471-4922. doi: 10.1016/j.pt.2014.10.001.

[317] N. G. Davies, S. Flasche, M. Jit, and K. E. Atkins. Modeling the effect of vaccination on selection for antibiotic resistance in Streptococcus pneumoniae. *Science Translational Medicine*, 13(606):eaaz8690, 2021. doi: doi:10.1126/scitranslmed.aaz8690. URL `https://www.science.org/doi/abs/10.1126/scitranslmed.aaz8690`.

[318] L. A. Hicks, L. H. Harrison, B. Flannery, J. L. Hadler, W. Schaffner, A. S. Craig, D. Jackson, A. Thomas, B. Beall, R. Lynfield, A. Reingold, M. M. Farley, and C. G. Whitney. Incidence of pneumococcal disease due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998-2004. *J Infect Dis*, 196(9):1346–54, 2007. doi: 10.1086/521626.

[319] D. Lehmann, J. Willis, H. C. Moore, C. Giele, D. Murphy, A. D. Keil, C. Harrison, K. Bayley, M. Watson, P. Richmond, and Vaccine Impact Surveillance Network. The changing epidemiology of invasive pneumococcal disease in Aboriginal and Non-Aboriginal Western Australians from 1997 through 2007 and emergence of nonvaccine serotypes. *Clin Infect Dis*, 50(11):1477–1486, 2010. ISSN 1058-4838. doi: 10.1086/652440. URL `http://dx.doi.org/10.1086/652440`.

[320] E. Y. Cho, H. M. Kang, J. Lee, J. H. Kang, E. H. Choi, and H. J. Lee. Changes in serotype distribution and antibiotic resistance of nasopharyngeal isolates of Streptococcus pneumoniae from children in Korea, after optional use of the 7-valent conjugate vaccine. *J Korean Med Sci*, 27(7):716–22, 2012. doi: 10.3346/jkms.2012.27.7.716. URL `https://www.ncbi.nlm.nih.gov/pubmed/22787364`.

[321] K. C. Thoon, C. Y. Chong, and N. W. Tee. Early impact of pneumococcal conjugate vaccine on invasive pneumococcal disease in Singapore children, 2005 through 2010. *Int J Infect Dis*, 16(3):e209–15, 2012. ISSN 1878-3511 (Electronic) 1201-9712 (Linking). doi: 10.1016/j.ijid.2011.11.014. URL `https://www.ncbi.nlm.nih.gov/pubmed/22281191`.

[322] M. R. Moore, R. Link-Gelles, W. Schaffner, R. Lynfield, C. Lexau, N. M. Bennett, S. Petit, S. M. Zansky, L. H. Harrison, A. Reingold, L. Miller, K. Scherzinger, A. Thomas, M. M. Farley, E. R. Zell, T. H. Taylor, T. Pondo, L. Rodgers, L. McGee, B. Beall, J. H. Jorgensen, and C. G. Whitney. Effect of use of 13-valent pneumococcal conjugate vaccine in children on invasive pneumococcal disease in children and adults in the USA: Analysis of multisite, population-based surveillance. *Lancet Infect Dis*, 15(3):301–309, 2015. ISSN 1473-3099. doi: http://dx.doi.org/10.1016/S1473-3099\%2814\%2971081-3. URL `http://www.journals.elsevier.com/the-lancet-infectious-diseases`.

[323] J. A. Lewnard and W. P. Hanage. Making sense of differences in pneumococcal serotype replacement. *Lancet Infect Dis*, 2019. ISSN 1473-3099. doi: https://doi.org/10.1016/S1473-3099(18)30660-1. URL `http://www.sciencedirect.com/science/article/pii/S1473309918306601`.

[324] M. Tin Tin Htar, D. Christopoulou, and H. J. Schmitt. Pneumococcal serotype evolution in western europe. *BMC Infectious Diseases*, 15:419, 2015. ISSN 1471-2334. doi: 10.1186/s12879-015-1147-x. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4606906/`.

[325] N. Ouldali, C. Levy, E. Varon, S. Bonacorsi, S. Bechet, R. Cohen, F. Angoulvant, and Network French Pediatric Meningitis. Incidence of paediatric pneumococcal meningitis and emergence of new serotypes: a time-series analysis of a 16-year french national survey. *Lancet Infect Dis*, 18(9):983–991, 2018. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(18)30349-9. URL https://www.ncbi.nlm.nih.gov/pubmed/30049623.

[326] C. Levy, N. Ouldali, L. Caeymaex, F. Angoulvant, E. Varon, and R. Cohen. Diversity of serotype replacement after pneumococcal conjugate vaccine implementation in Europe. *J Pediatr*, 213:252–253 e3, 2019. ISSN 1097-6833 (Electronic) 0022-3476 (Linking). doi: 10.1016/j.jpeds.2019.07.057. URL https://www.ncbi.nlm.nih.gov/pubmed/31561776.

[327] K. L. O'Brien, E. V. Millar, E. R. Zell, M. Bronsdon, R. Weatherholtz, R. Reid, J. Becenti, S. Kvamme, C. G. Whitney, and M. Santosham. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. *J Infect Dis*, 196(8):1211–20, 2007. doi: 10.1086/521833. URL https://www.ncbi.nlm.nih.gov/pubmed/17955440.

[328] W. P. Hausdorff, M. K. Van Dyke, and T. Van Effelterre. Serotype replacement after pneumococcal vaccination. *Lancet*, 379(9824):1387–1388, 2011. ISSN 0140-6736. doi: 10.1016/S0140-6736(12)60589-3. URL http://dx.doi.org/10.1016/S0140-6736(12)60589-3.

[329] C. Savulescu, P. Krizova, A. Lepoutre, J. Mereckiene, D. F. Vestrheim, P. Ciruela, M. Ordobas, M. Guevara, E. McDonald, E. Morfeldt, J. Kozakova, E. Varon, S. Cotter, B. A. Winje, C. Munoz-Almagro, L. Garcia, J. Castilla, A. Smith, B. Henriques-Normark, L. P. Celentano, and G. Hanquet. Effect of high-valency pneumococcal conjugate vaccines on invasive pneumococcal disease in children in SpIDnet countries: an observational multicentre study. *The Lancet Respiratory Medicine*, 5(8):648–656, 2017. ISSN 2213-2600. doi: https://doi.org/10.1016/S2213-2600(17)30110-8. URL http://www.sciencedirect.com/science/article/pii/S2213260017301108.

[330] M. Corcoran, I. Vickers, J. Mereckiene, S. Murchan, S. Cotter, M. Fitzgerald, M. McElligott, M. Cafferkey, D. O'Flanagan, R. Cunney, and H. Humphreys. The epidemiology of invasive pneumococcal disease in older adults in the post-PCV era. has there been a herd effect? *Epidemiol Infect*, 145(11):2390–2399, 2017. ISSN 0950-2688. doi: 10.1017/s0950268817001194.

[331] C. Colijn, J. Corander, and N. J. Croucher. Designing ecologically optimized pneumococcal vaccines using population genomics. *Nature Microbiology*, 5(3):473–485, 2020. ISSN 2058-5276. doi: 10.1038/s41564-019-0651-y. URL https://doi.org/10.1038/s41564-019-0651-y.

[332] M. Corcoran, J. Mereckiene, S. Cotter, S. Murchan, R. Cunney, and H. Humphreys. Invasive Streptococcus pneumoniae infections and vaccine failures in children in Ireland from the postvaccine era from 2007 to 2018. *Pediatr Infect Dis J*, Online First, 2019. ISSN 0891-3668. doi: 10.1097/inf.0000000000002549. URL https://journals.lww.com/pidj/Fulltext/onlinefirst/Invasive_Streptococcus_pneumoniae_Infections_and.96312.aspx.

[333] E. Balsells, L. Guillot, H. Nair, and M.H. Kyaw. Serotype distribution of Streptococcus pneumoniae causing invasive disease in children in the post-PCV era: A systematic review and meta-analysis. *PloS One*, 12(5):e0177113, 2017. ISSN 1932-6203. doi: https://dx.doi.org/10.1371/journal.pone.0177113. URL http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=28486544.

[334] The World Bank. High income, 2019. URL https://data.worldbank.org/income-level/high-income.

[335] S. A. Nzenze, A. von Gottberg, T. Shiri, N. van Niekerk, L. de Gouveia, A. Violari, M. C. Nunes, and S. A. Madhi. Temporal changes in pneumococcal colonization in HIV-infected and HIV-uninfected mother-child pairs following transitioning from 7-valent to 13-valent pneumococcal conjugate vaccine, Soweto, South Africa. *The Journal of*

Infectious Diseases, 212(7):1082–1092, 2015. ISSN 0022-1899. doi: 10.1093/infdis/jiv167. URL https://doi.org/10.1093/infdis/jiv167.

[336] H. Grundmann, S. Hori, and G. Tanner. Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. *Journal of Clinical Microbiology*, 39(11):4190–4192, 2001. ISSN 0095-1137 1098-660X. doi: 10.1128/JCM.39.11.4190-4192.2001. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC88515/.

[337] SY. Chen, Z. Feng, and X. Yi. A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease*, 9(6):1725–1729, 2017. doi: 10.21037/jtd.2017.05.34. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5506159/.

[338] J. W. Eyre and J. W. Washbourn. Varieties and virulence of the pneumococcus. *The Lancet*, 153(3932):19–22, 1899. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(01)78950-7. URL https://www.sciencedirect.com/science/article/pii/S0140673601789507.

[339] D. E. Briles, M. J. Crain, B. M. Gray, C. Forman, and J. Yother. Strong association between capsular type and virulence for mice among human isolates of streptococcus pneumoniae. *Infect Immun*, 60(1):111–6, 1992. ISSN 0019-9567 (Print) 0019-9567 (Linking). doi: 10.1128/iai.60.1.111-116.1992. URL https://www.ncbi.nlm.nih.gov/pubmed/1729176.

[340] J. A. Lewnard, N. Givon-Lavi, D. M. Weinberger, M. Lipsitch, and R. Dagan. Pan-serotype reduction in progression of Streptococcus pneumoniae to otitis media after rollout of pneumococcal conjugate vaccines. *Clin Infect Dis*, 65(11):1853–1861, 2017. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1093/cid/cix673. URL https://www.ncbi.nlm.nih.gov/pubmed/29020218.

[341] J. Y. Song, M. H. Nahm, and M. A. Moseley. Clinical implications of pneumococcal serotypes: invasive disease potential, clinical presentations, and antibiotic resistance. *J Korean Med Sci*, 28(1):4–15, 2013. ISSN 1598-6357 (Electronic) 1011-8934 (Linking).

doi: 10.3346/jkms.2013.28.1.4. URL https://www.ncbi.nlm.nih.gov/pubmed/23341706.

[342] T. Smith, D. Lehmann, J. Montgomery, M. Gratten, I. D. Riley, and M. P. Alpers. Acquisition and invasiveness of different serotypes of Streptococcus pneumoniae in young children. *Epidemiol Infect*, 111(1):27–39, 1993. URL https://www.ncbi.nlm.nih.gov/pubmed/8348930.

[343] S. W. Lo, R. A. Gladstone, A. J. van Tonder, J. A. Lees, M. du Plessis, R. Benisty, N. Givon-Lavi, P. A. Hawkins, J. E. Cornick, B. Kwambana-Adams, P. Y. Law, P. L. Ho, M. Antonio, D. B. Everett, R. Dagan, A. von Gottberg, K. P. Klugman, L. McGee, R. F. Breiman, S. D. Bentley, and Consortium Global Pneumococcal Sequencing. Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-pcv13 era: an international whole-genome sequencing study. *Lancet Infect Dis*, 19(7):759–769, 2019. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(19)30297-X. URL https://www.ncbi.nlm.nih.gov/pubmed/31196809.

[344] E. Balsells, R. Dagan, I. Yildirim, P. P. Gounder, A. Steens, C. Munoz-Almagro, C. Mameli, R. Kandasamy, N. Givon Lavi, L. Daprai, A. van der Ende, K. Trzcinski, S. A. Nzenze, S. Meiring, D. Foster, L. R. Bulkow, K. Rudolph, A. Valero-Rello, S. Ducker, D. F. Vestrheim, A. von Gottberg, S. I. Pelton, G. Zuccotti, A. J. Pollard, E. A. M. Sanders, H. Campbell, S. A. Madhi, H. Nair, and M. H. Kyaw. The relative invasive disease potential of Streptococcus pneumoniae among children after PCV introduction: A systematic review and meta-analysis. *J Infect*, 77(5):368–378, 2018. ISSN 0163-4453. doi: 10.1016/j.jinf.2018.06.004.

[345] J. Southern, N. Andrews, P. Sandu, C. L. Sheppard, P. A. Waight, N. K. Fry, A. J. Van Hoek, and E. Miller. Pneumococcal carriage in children and their household contacts six years after introduction of the 13-valent pneumococcal conjugate vaccine in England. *PLoS One*, 13(5):e0195799, 2018. ISSN 1932-6203 (Electronic) 1932-6203

(Linking). doi: 10.1371/journal.pone.0195799. URL `https://www.ncbi.nlm.nih.gov/pubmed/29799839`.

[346] A. Zwahlen, J. A. Winkelstein, and E. R. Moxon. Surface determinants of Haemophilus influenzae pathogenicity: Comparative virulence of capsular transformants in normal and complement-depleted rats. *The Journal of Infectious Diseases*, 148(3):385–394, 1983. ISSN 0022-1899. doi: 10.1093/infdis/148.3.385. URL `https://doi.org/10.1093/infdis/148.3.385`.

[347] T. Kelly, J. P. Dillard, and J. Yother. Effect of genetic switching of capsular type on virulence of Streptococcus pneumoniae. *Infection and Immunity*, 62(5):1813–1819, 1994. URL `https://iai.asm.org/content/iai/62/5/1813.full.pdf`.

[348] L. J. Hathaway, D. Grandgirard, L. G. Valente, M. G. Täuber, and S. L. Leib. Streptococcus pneumoniae capsule determines disease severity in experimental pneumococcal meningitis. *Open biology*, 6(3):150269, 2016. ISSN 2046-2441. doi: 10.1098/rsob.150269. URL `https://pubmed.ncbi.nlm.nih.gov/27009189https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4821241/`.

[349] F. Z. Hu, R. Eutsey, A. Ahmed, N. Frazao, E. Powell, N. L. Hiller, T. Hillman, F. J. Buchinsky, R. Boissy, B. Janto, J. Kress-Bennett, M. Longwell, S. Ezzo, J. C. Post, M. Nesin, A. Tomasz, and G. D. Ehrlich. In vivo capsular switch in Streptococcus pneumoniae - analysis by whole genome sequencing. *PloS one*, 7(11):e47983–e47983, 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0047983. URL `https://pubmed.ncbi.nlm.nih.gov/23144841https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3493582/`.

[350] Y. Mizrachi Nebenzahl, N. Porat, S. Lifshitz, S. Novick, A. Levi, E. Ling, O. Liron, S. Mordechai, R.K. Sahu, and R. Dagan. Virulence of Streptococcus pneumoniae may be determined independently of capsular polysaccharide. *FEMS Microbiology Letters*, 233(1):147–152, 2004. ISSN 0378-1097. doi: 10.1016/j.femsle.2004.02.003. URL `https://doi.org/10.1016/j.femsle.2004.02.003`.

[351] R. A. Gladstone, S. W. Lo, J. A. Lees, N. J. Croucher, A. J. van Tonder, J. Corander, A. J. Page, P. Marttinen, L. J. Bentley, T. J. Ochoa, P. L. Ho, M. du Plessis, J. E. Cornick, B. Kwambana-Adams, R. Benisty, S. A. Nzenze, S. A. Madhi, P. A. Hawkins, D. B. Everett, M. Antonio, R. Dagan, K. P. Klugman, A. von Gottberg, L. McGee, R. F. Breiman, S. D. Bentley, and Global Pneumococcal Sequencing Consortium. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine*, 43:338–346, 2019. ISSN 2352-3964. doi: 10.1016/j.ebiom. 2019.04.021. URL `https://www.ncbi.nlm.nih.gov/pubmed/31003929https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6557916/`.

[352] R. Sá-Leão, F. Pinto, S. Aguiar, S. Nunes, J. A. Carriço, N. Frazão, N. Gonçalves-Sousa, J. Melo-Cristino, H. de Lencastre, and M. Ramirez. Analysis of invasiveness of pneumococcal serotypes and clones circulating in Portugal before widespread use of conjugate vaccines reveals heterogeneous behavior of clones expressing the same serotype. *Journal of Clinical Microbiology*, 49(4):1369–1375, 2011. doi: 10.1128/jcm. 01763-10. URL `https://jcm.asm.org/content/jcm/49/4/1369.full.pdf`.

[353] C. Donati, N. L. Hiller, H. Tettelin, A. Muzzi, N. J. Croucher, S. V. Angiuoli, M. Oggioni, J. C. Dunning Hotopp, F. Z. Hu, D. R. Riley, A. Covacci, T. J. Mitchell, S. D. Bentley, M. Kilian, G. D. Ehrlich, R. Rappuoli, E. R. Moxon, and V. Masignani. Structure and dynamics of the pan-genome of Streptococcus pneumoniae and closely related species. *Genome biology*, 11(10):R107–R107, 2010. doi: 10.1186/gb-2010-11-10-r107. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3218663/`.

[354] C. Obert, J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. I. Tuomanen, and C. J. Orihuela. Identification of a candidate Streptococcus pneumoniae core genome and regions of diversity correlated with invasive pneumococcal disease. *Infection and immunity*, 74(8):4766–4777, 2006. doi: 10.1128/IAI.00316-06. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1539573/`.

[355] J. A. Lees, B. Ferwerda, P. H. C. Kremer, N. E. Wheeler, M. V. Serón, N. J. Croucher, R. A. Gladstone, H. J. Bootsma, N. Y. Rots, A. J. Wijmega-Monsuur, E. A. M. Sanders,

K. Trzciński, A. L. Wyllie, A. H. Zwinderman, L. H. van den Berg, W. van Rheenen, J. H. Veldink, Z. B. Harboe, L. F. Lundbo, L. C. P. G. M. de Groot, N. M. van Schoor, N. van der Velde, L. H. Ängquist, T. I. A. Sørensen, E. A. Nohr, A. J. Mentzer, T. C. Mills, J. C. Knight, M. du Plessis, S. Nzenze, J. N. Weiser, J. Parkhill, S. Madhi, T. Benfield, A. von Gottberg, A. van der Ende, M. C. Brouwer, J. C. Barrett, S. D. Bentley, and D. van de Beek. Joint sequencing of human and pathogen genomes reveals the genetics of pneumococcal meningitis. *Nature communications*, 10(1):2176–2176, 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-09976-3. URL `https://www.ncbi.nlm.nih.gov/pubmed/31092817https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6520353/`.

[356] M. Nurhonen and K. Auranen. Optimal serotype compositions for pneumococcal conjugate vaccination under serotype replacement. *PLOS Computational Biology*, 10 (2):e1003477, 2014. doi: 10.1371/journal.pcbi.1003477. URL `https://doi.org/10.1371/journal.pcbi.1003477`.

[357] A. Løchen, 2020. URL `https://github.com/alochen/pneumo-invasiveness`.

[358] A. K. Takala, J. Vuopio-Varkila, E. Tarkka, M. Leinonen, and J. M. Musser. Subtyping of common pediatric pneumococcal serotypes from invasive disease and pharyngeal carriage in Finland. *J Infect Dis*, 173(1):128–35, 1996.

[359] C. Azzari, M. Cortimiglia, F. Nieddu, M. Moriondo, G. Indolfi, R. Mattei, M. Zuliani, B. Adriani, R. Degl'Innocenti, G. Consales, D. Aquilini, G. Bini, M. E. Di Natale, C. Canessa, S. Ricci, E. de Vitis, G. Mangone, A. Bechini, P. Bonanni, A. Pasinato, and M. Resti. Pneumococcal serotype distribution in adults with invasive disease and in carrier children in Italy: Should we expect herd protection of adults through infants' vaccination? *Hum Vaccin Immunother*, 12(2):344–50, 2016. ISSN 2164-5515. doi: 10.1080/21645515.2015.1102811.

[360] X. Valles, A. Roca, F. Lozano, L. Morais, B. Suarez, F. Casals, I. Mandomando, B. Sigauque, D. Nhalungo, C. Esquinas, L. Quinto, P. L. Alonso, and A. Torres. Serotype-

specific pneumococcal disease may be influenced by mannose-binding lectin deficiency. *Eur Respir J*, 36(4):856–63, 2010. ISSN 0903-1936. doi: 10.1183/09031936.00171409.

[361] N. Principi and P. Marchisio. Epidemiology of Streptococcus pneumoniae in italian children. *Acta Paediatr Suppl*, 89(435):40–3, 2000.

[362] L. E. Smart, D. J. Platt, and M. C. Timbury. A comparison of the distribution of pneumococcal types in systemic disease and the upper respiratory tract in adults and children. *Epidemiol Infect*, 98(2):203–9, 1987.

[363] I. Galanis, A. Lindstrand, J. Darenberg, S. Browall, P. Nannapaneni, K. Sjöström, E. Morfeldt, P. Naucler, M. Blennow, Å. Örtqvist, and B. Henriques-Normark. Effects of pcv7 and pcv13 on invasive pneumococcal disease and carriage in stockholm, sweden. *European Respiratory Journal*, 47(4):1208–1218, 2016. doi: 10.1183/13993003.01451-2015. URL https://erj.ersjournals.com/content/erj/47/4/1208.full.pdf.

[364] A. R. Cullotta, H. D. Kalter, J. Delgado, R. H. Gilman, R. R. Facklam, B. Velapatino, J. Coronel, L. Cabrera, and M. Urbina. Antimicrobial susceptibilities and serotype distribution of Streptococcus pneumoniae isolates from a low socioeconomic area in lima, peru. *Clin Diagn Lab Immunol*, 9(6):1328–31, 2002.

[365] M. C. Mogdasy, T. Camou, C. Fajardo, and M. Hortal. Colonizing and invasive strains of Streptococcus pneumoniae in Uruguayan children: type distribution and patterns of antibiotic resistance. *Pediatr Infect Dis J*, 11(8):648–52, 1992.

[366] J. A. Scott, A. J. Hall, A. Hannington, R. Edwards, S. Mwarumba, B. Lowe, D. Griffiths, D. Crook, and K. Marsh. Serotype distribution and prevalence of resistance to benzylpenicillin in three representative populations of Streptococcus pneumoniae isolates from the coast of Kenya. *Clin Infect Dis*, 27(6):1442–50, 1998.

[367] S. Browall, E. Backhaus, P. Naucler, I. Galanis, K. Sjostrom, D. Karlsson, S. Berg, J. Luthander, M. Eriksson, C. Spindler, M. Ejdeback, B. Trollfors, J. Darenberg, M. Kalin, A. Ortqvist, R. Andersson, and B. Henriques-Normark. Clinical manifestations of

invasive pneumococcal disease by vaccine and non-vaccine types. *Eur Respir J*, 44(6): 1646–57, 2014. ISSN 0903-1936. doi: 10.1183/09031936.00080814.

[368] S. Levidiotou, G. Vrioni, G. Tzanakaki, C. Pappa, H. Gesouli, C. Gartzonika, D. Papamichail, and J. Kremastinou. Serotype distribution of Streptococcus pneumoniae in north-western Greece and implications for a vaccination programme. *FEMS Immunol Med Microbiol*, 48(2):179–82, 2006. doi: 10.1111/j.1574-695X.2006.00126.x.

[369] W. P. Hanage, T. H. Kaijalainen, R. K. Syrjanen, K. Auranen, M. Leinonen, P. H. Makela, and B. G. Spratt. Invasiveness of serotypes and clones of Streptococcus pneumoniae among children in Finland. *Infect Immun*, 73(1):431–5, 2005. doi: 10.1128/IAI.73.1. 431-435.2005. URL https://www.ncbi.nlm.nih.gov/pubmed/15618181.

[370] C. Mameli, V. Fabiano, L. Daprai, G. Bedogni, M. Faccini, M. L. Garlaschi, F. Penagini, D. Dilillo, E. Torresani, M. Gramegna, and G. V. Zuccotti. A longitudinal study of Streptococcus pneumoniae carriage in healthy children in the 13-valent pneumococcal conjugate vaccine era. *Hum Vaccin Immunother*, 11(4):811–7, 2015. ISSN 2164-554X (Electronic) 2164-5515 (Linking). doi: 10.1080/21645515.2015.1010945. URL https://www.ncbi.nlm.nih.gov/pubmed/25751237.

[371] A. Lovlie, D. F. Vestrheim, I. S. Aaberge, and A. Steens. Changes in pneumococcal carriage prevalence and factors associated with carriage in Norwegian children, four years after introduction of pcv13. *BMC Infect Dis*, 20(1):29, 2020. ISSN 1471-2334 (Electronic) 1471-2334 (Linking). doi: 10.1186/s12879-019-4754-0. URL https://www.ncbi.nlm.nih.gov/pubmed/31924177.

[372] E. del Amo, P. Brotons, M. Monsonis, M. Trivino, M. Inigo, L. Selva, R. Sa-Leao, and C. Munoz-Almagro. High invasiveness of pneumococcal serotypes included in the new generation of conjugate vaccines. *Clin Microbiol Infect*, 20(7):684–9, 2014. ISSN 1198-743x. doi: 10.1111/1469-0691.12422.

[373] B. M. Gray, 3rd Converse, G. M., and Jr. Dillon, H. C. Serotypes of Streptococcus pneumoniae causing disease. *J Infect Dis*, 140(6):979–83, 1979.

[374] I. Jroundi, C. Mahraoui, R. Benmessaoud, C. Moraleda, C. Munoz Almagro, M. Seffar, H. Tligui, S. C. Kettani, B. S. Benjelloun, P. L. Alonso, and Q. Bassat. Streptococcus pneumoniae carriage among healthy and sick pediatric patients before the generalized implementation of the 13-valent pneumococcal vaccine in Morocco from 2010 to 2011. *J Infect Public Health*, 10(2):165–170, 2017. ISSN 1876-0341. doi: 10.1016/j.jiph.2016.02.012.

[375] J. D. Kellner, A. McGeer, M. S. Cetron, D. E. Low, J. C. Butler, A. Matlow, J. Talbot, and E. L. Ford-Jones. The use of Streptococcus pneumoniae nasopharyngeal isolates from healthy children to predict features of invasive disease. *Pediatr Infect Dis J*, 17(4):279–86, 1998.

[376] E. L. Parra, F. De La Hoz, P. L. Diaz, O. Sanabria, M. E. Realpe, and J. Moreno. Changes in Streptococcus pneumoniae serotype distribution in invasive disease and nasopharyngeal carriage after the heptavalent pneumococcal conjugate vaccine introduction in Bogota, Colombia. *Vaccine*, 31(37):4033–8, 2013. ISSN 0264-410x. doi: 10.1016/j.vaccine.2013.04.074.

[377] I. A. Rivera-Olivero, B. del Nogal, M. C. Sisco, D. Bogaert, P. W. Hermans, and J. H. de Waard. Carriage and invasive isolates of Streptococcus pneumoniae in Caracas, Venezuela: the relative invasiveness of serotypes and vaccine coverage. *Eur J Clin Microbiol Infect Dis*, 30(12):1489–95, 2011. ISSN 0934-9723. doi: 10.1007/s10096-011-1247-5.

[378] A. Sandgren, K. Sjostrom, B. Olsson-Liljequist, B. Christensson, A. Samuelsson, G. Kronvall, and B. Henriques Normark. Effect of clonal and serotype-specific properties on the invasive capacity of Streptococcus pneumoniae. *J Infect Dis*, 189(5):785–96, 2004. doi: 10.1086/381686.

[379] J. R. Scott, E. V. Millar, M. Lipsitch, L. H. Moulton, R. Weatherholtz, M. J. Perilla, D. M. Jackson, B. Beall, M. J. Craig, R. Reid, M. Santosham, and K. L. O'Brien. Impact of more than a decade of pneumococcal conjugate vaccine use on carriage and invasive potential in Native American communities. *J Infect Dis*, 205(2):280–8, 2012. ISSN 0022-1899. doi: 10.1093/infdis/jir730.

[380] C. L. Trotter, P. Waight, N. J. Andrews, M. Slack, A. Efstratiou, R. George, and E. Miller. Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. *J Infect*, 60(3):200–8, 2010. ISSN 1532-2742 (Electronic) 0163-4453 (Linking). doi: 10.1016/j.jinf.2009.12.008. URL `https://www.ncbi.nlm.nih.gov/pubmed/20035785`.

[381] E. Varon, R. Cohen, S. Bechet, C. Doit, and C. Levy. Invasive disease potential of pneumococci before and after the 13-valent pneumococcal conjugate vaccine implementation in children. *Vaccine*, 33(46):6178–85, 2015. ISSN 0264-410x. doi: 10.1016/j.vaccine.2015.10.015.

[382] I. Yildirim, W. P. Hanage, M. Lipsitch, K. M. Shea, A. Stevenson, J. Finkelstein, S. S. Huang, G. M. Lee, K. Kleinman, and S. I. Pelton. Serotype specific invasive capacity and persistent reduction in invasive pneumococcal disease. *Vaccine*, 29(2):283–8, 2010. doi: 10.1016/j.vaccine.2010.10.032. URL `https://www.ncbi.nlm.nih.gov/pubmed/21029807`.

[383] H. Zemlickova, V. Jakubu, P. Urbaskova, J. Motlova, M. Musilek, and V. Adamkova. Serotype-specific invasive disease potential of Streptococcus pneumoniae in Czech children. *J Med Microbiol*, 59(Pt 9):1079–83, 2010. ISSN 0022-2615. doi: 10.1099/jmm.0.018390-0.

[384] M. Vissers, A. J. Wijmenga-Monsuur, M. J. Knol, P. Badoux, M. A. van Houten, A. van der Ende, E. A. M. Sanders, and N. Y. Rots. Increased carriage of non-vaccine serotypes with low invasive disease potential four years after switching to the 10-valent pneumococcal conjugate vaccine in the Netherlands. *PLoS One*, 13(3):e0194823, 2018. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0194823. URL `https://www.ncbi.nlm.nih.gov/pubmed/29601605`.

[385] M. Auzenbergs, C. Correia-Gomes, T. Economou, R. Lowe, and K. M. O'Reilly. Desirable bugs in models of infectious diseases. *Epidemics*, 29:100361, 2019. ISSN 1878-0067

(Electronic) 1878-0067 (Linking). doi: 10.1016/j.epidem.2019.100361. URL `https://www.ncbi.nlm.nih.gov/pubmed/31668494`.

[386] J. K. Kruschke and T. M. Liddell. The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon Bull Rev*, 25(1):178–206, 2018. ISSN 1531-5320 (Electronic) 1069-9384 (Linking). doi: 10.3758/s13423-016-1221-4. URL `https://www.ncbi.nlm.nih.gov/pubmed/28176294`.

[387] A. Arguedas, K. Trzciński, K.L. O'Brien, D.M. Ferreira, A.L. Wyllie, D. Weinberger, L. Danon, S.I. Pelton, C. Azzari, L.L. Hammitt, R. Sá-Leão, M.C. Brandileone, S. Saha, J. Suaya, R. Isturiz, L. Jodar, and B.D. Gessner. Upper respiratory tract colonization with Streptococcus pneumoniae in adults. *Expert Review of Vaccines*, 19(4):353–366, 2020. ISSN 1476-0584. doi: 10.1080/14760584.2020.1750378. URL `https://doi.org/10.1080/14760584.2020.1750378`.

[388] D. Makowski, M. S. Ben-Shachar, and D. Lüdecke. bayestestr: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, 4(40):1541, 2019. doi: 10.21105/joss.01541. URL `https://doi.org/10.21105/joss.01541`.

[389] H. Jeffreys. *Theory of probability*. The International series of monographs on physics. Clarendon Press, Oxford,, 3d edition, 1961.

[390] C. Gibson. Population of the 100 largest cities and other urban places in the United States: 1790 to 1990, 1998. URL `https://www.census.gov/library/working-papers/1998/demo/POP-twps0027.html`.

[391] Instituto Nacional de Estadistica Venezuela. Censo nacional de poblacion y vivienda, 2011. URL `http://www.redatam.ine.gob.ve/Censo2011/index.html`.

[392] Czech Statistical Office. Age distribution of the population of the Czech Republic - 2005, 2005. URL `https://www.czso.cz/csu/czso/4003-06-2005-01n`.

[393] Papua New Guinea National Statistical Office. National population and housing census ward population profile: Highlands region, 2011. URL `https://www.nso.gov.pg/index.php/population-and-social/other-indicators#highlands-region`.

[394] Morocco Haut-Commissariat au Plan. Annuaire statistique de la région Rabat-Salé-Zemmour-Zaër, 2010, 2012. URL `https://www.hcp.ma/downloads/Annuaires-statistiques-regionaux_t11956.html`.

[395] United Nations Department of Economic and Social Affairs. Population by age groups - both sexes, 2020. URL `https://population.un.org/wpp/Download/Standard/Population/`.

[396] Statistics Canada. Data tables, 1996 Canada, 1996. URL `https://www.statcan.gc.ca/https://www12.statcan.gc.ca/English/census96/data/tables/Rp-eng.cfm?LANG=E&APATH=3&DETAIL=1&DIM=0&FL=A&FREE=1&GC=0&GID=0&GK=0&GRP=1&PID=1029&PRID=0&PTYPE=89103&S=0&SHOWALL=No&SUB=0&Temporal=2006&THEME=1&VID=0&VNAMEE=&VNAMEF=`.

[397] Instituto Nacional de Estadistica Portugal. Censos 2001, 2002.

[398] Statistics Sweden. Population by region, marital status, age and sex. year 1968 - 2018, 2018. URL `http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101A/BefolkningNy/?rxid=e550e75b-4cef-4e77-b5e5-83adcc09c97e`.

[399] Institut National de la Statistique et des Etudes Economiques (INSEE). Age pyramid 2020 – France and metropolitan France, 2008. URL `https://www.insee.fr/en/statistiques/3312960?champ=fe&lang=en&annee=2018`.

[400] Navajo Epidemiology Center Navajo Division of Health. Navajo Population Profile 2010 U.S. Census, 2013. URL `https://www.nec.navajo-nsn.gov/Portals/0/Reports/NN2010PopulationProfile.pdf`.

[401] M. Zancolli, P. Canepa, A. Ceravolo, V. Parodi, and F. Ansaldi. Determinants of invasiveness and ability to cause invasive pneumococcal disease, pneumonia and acute otitis media of different serotypes of Streptococcus pneumoniae. *J Prev Med Hyg*, 52(2): 47–54, 2011.

[402] R. Weatherholtz, E. V. Millar, L. H. Moulton, R. Reid, K. Rudolph, M. Santosham, and K. L. O'Brien. Invasive pneumococcal disease a decade after pneumococcal conjugate vaccine use in an American Indian population at high risk for disease. *Clin Infect Dis*, 50(9):1238–46, 2010. ISSN 1058-4838. doi: 10.1086/651680.

[403] H. M. Garcia Garrido, A. M. R. Mak, Fwnm Wit, G. W. M. Wong, M. J. Knol, A. Vollaard, M. W. T. Tanck, A. Van Der Ende, M. P. Grobusch, and A. Goorhuis. Incidence and risk factors for invasive pneumococcal disease and community-acquired pneumonia in human immunodeficiency virus-infected individuals in a high-income setting. *Clin Infect Dis*, 71(1):41–50, 2020. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1093/cid/ciz728. URL https://www.ncbi.nlm.nih.gov/pubmed/31634398.

[404] A. Torres, F. Blasi, N. Dartois, and M. Akova. Which individuals are at increased risk of pneumococcal disease and why? Impact of COPD, asthma, smoking, diabetes, and/or chronic heart disease on community-acquired pneumonia and invasive pneumococcal disease. *Thorax*, 70(10):984–9, 2015. ISSN 1468-3296 (Electronic) 0040-6376 (Linking). doi: 10.1136/thoraxjnl-2015-206780. URL https://www.ncbi.nlm.nih.gov/pubmed/26219979.

[405] A. Wong, T. J. Marrie, S. Garg, J. D. Kellner, G. J. Tyrrell, and Spat Group. Increased risk of invasive pneumococcal disease in haematological and solid-organ malignancies. *Epidemiol Infect*, 138(12):1804–10, 2010. ISSN 1469-4409 (Electronic) 0950-2688 (Linking). doi: 10.1017/S0950268810000919. URL https://www.ncbi.nlm.nih.gov/pubmed/20429967.

[406] P. E. Kim, D. M. Musher, W. P. Glezen, M. C. Rodriguez-Barradas, W. K. Nahm, and C. E. Wright. Association of invasive pneumococcal disease with season, atmospheric

conditions, air pollution, and the isolation of respiratory viruses. *Clin Infect Dis*, 22 (1):100–6, 1996. doi: 10.1093/clinids/22.1.100. URL https://www.ncbi.nlm.nih.gov/pubmed/8824973.

[407] A. A. Palmu, P. De Wals, M. Toropainen, S. N. Ladhani, G. Deceuninck, M. J. Knol, E. A. M. Sanders, and E. Miller. Similar impact and replacement disease after pneumococcal conjugate vaccine introduction in hospitalised children with invasive pneumococcal disease in Europe and North America. *Vaccine*, 39(11):1551–1555, 2021. doi: 10.1016/j.vaccine.2021.01.070. URL https://www.ncbi.nlm.nih.gov/pubmed/33610373.

[408] Jonathan A. McCullers. The co-pathogenesis of influenza viruses with bacteria in the lung. *Nature Reviews Microbiology*, 12:252, 2014. doi: 10.1038/nrmicro3231. URL https://doi.org/10.1038/nrmicro3231.

[409] N. Wolter, S. Tempia, C. Cohen, S. A. Madhi, M. Venter, J. Moyes, S. Walaza, B. Malope-Kgokong, M. Groome, M. du Plessis, V. Magomani, M. Pretorius, O. Hellferscee, H. Dawood, K. Kahn, E. Variava, K. P. Klugman, and A. von Gottberg. High nasopharyngeal pneumococcal density, increased by viral coinfection, is associated with invasive pneumococcal pneumonia. *J Infect Dis*, 210(10):1649–57, 2014. ISSN 0022-1899. doi: 10.1093/infdis/jiu326.

[410] M. T. Phillips, J. L. Warren, N. Givon-Lavi, A. Tothpal, G. Regev-Yochay, R. Dagan, and D. M. Weinberger. Evaluating post-vaccine expansion patterns of pneumococcal serotypes. *Vaccine*, 38(49):7756–7763, 2020. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2020.10.045. URL https://www.ncbi.nlm.nih.gov/pubmed/33164799.

[411] G. Regev-Yochay, M. Raz, R. Dagan, N. Porat, B. Shainberg, E. Pinco, N. Keller, and E. Rubinstein. Nasopharyngeal carriage of Streptococcus pneumoniae by adults and children in community and family settings. *Clin Infect Dis*, 38(5):632–9, 2004. ISSN

1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1086/381547. URL `https://www.ncbi.nlm.nih.gov/pubmed/14986245`.

[412] D. M. Weinberger, V. E. Pitzer, G. Regev-Yochay, N. Givon-Lavi, and R. Dagan. Association between the decline in pneumococcal disease in unimmunized adults and vaccine-derived protection against colonization in toddlers and preschool-aged children. *American Journal of Epidemiology*, 188(1):160–168, 2019. ISSN 1476-6256 0002-9262. doi: 10.1093/aje/kwy219. URL `https://pubmed.ncbi.nlm.nih.gov/30462150https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6321804/`.

[413] A. N. Horácio, C. Silva-Costa, J. P. Lopes, M. Ramirez, J. Melo-Cristino, and Portuguese Group for the Study of Streptococcal Infections. Serotype 3 remains the leading cause of invasive pneumococcal disease in adults in Portugal (2012–2014) despite continued reductions in other 13-valent conjugate vaccine serotypes. *Frontiers in Microbiology*, 7(1616), 2016. ISSN 1664-302X. doi: 10.3389/fmicb.2016.01616. URL `https://www.frontiersin.org/article/10.3389/fmicb.2016.01616`.

[414] A. Steens, M. A. R. Bergsaker, I. S. Aaberge, K. Rønning, and D. F. Vestrheim. Prompt effect of replacing the 7-valent pneumococcal conjugate vaccine with the 13-valent vaccine on the epidemiology of invasive pneumococcal disease in Norway. *Vaccine*, 31(52):6232–6238, 2013. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2013.10.032. URL `http://www.sciencedirect.com/science/article/pii/S0264410X13014084`.

[415] G. Béraud, S. Kazmerrciak, P. Beutels, D. Levy-Bruhl, X. Lenne, N. Mielcarek, Y. Yazdanpanah, P. Boëlle, N. Hens, and B. Dervaux. Social contact data for France, 2020. URL `https://doi.org/10.5281/zenodo.1157918`.

[416] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med*, 5(3):e74, 2008. ISSN 1549-1676 (Electronic) 1549-1277 (Linking). doi:

10.1371/journal.pmed.0050074. URL `https://www.ncbi.nlm.nih.gov/pubmed/18366252`.

[417] C. Turner, P. Turner, V. Carrara, K. Burgoine, S. Tha Ler Htoo, W. Watthanaworawit, N. P. Day, N. J. White, D. Goldblatt, and F. Nosten. High rates of pneumonia in children under two years of age in a South East Asian refugee population. *PLoS One*, 8(1):e54026, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0054026. URL `https://www.ncbi.nlm.nih.gov/pubmed/23320118`.

[418] P. Turner, C. Turner, A. Jankhot, K. Phakaudom, F. Nosten, and D. Goldblatt. Field evaluation of culture plus latex sweep serotyping for detection of multiple pneumococcal serotype colonisation in infants and young children. *PLOS ONE*, 8(7):e67933, 2013. doi: 10.1371/journal.pone.0067933. URL `https://doi.org/10.1371/journal.pone.0067933`.

[419] J. A. Lees, N. J. Croucher, D. Goldblatt, F. Nosten, J. Parkhill, C. Turner, P. Turner, and S. D. Bentley. Genome-wide identification of lineage and locus specific variation associated with pneumococcal carriage duration. *eLife*, 6:e26255, 2017. ISSN 2050-084X. doi: 10.7554/eLife.26255. URL `https://doi.org/10.7554/eLife.26255`.

[420] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125:306–312, 2013. ISSN 1538-3873. doi: 10.1086/670067. URL `http://dx.doi.org/10.1086/670067`.

[421] P. Caley and D. Ramsey. Estimating disease transmission in wildlife, with emphasis on leptospirosis and bovine tuberculosis in possums, and effects of fertility control. *Journal of Applied Ecology*, 38(6):1362–1370, 2001. ISSN 0021-8901. doi: https://doi.org/10.1046/j.0021-8901.2001.00676.x. URL `https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1046/j.0021-8901.2001.00676.x`.

[422] L. Willem, F. Verelst, J. Bilcke, N. Hens, and P. Beutels. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC infectious diseases*, 17(1):612–612, 2017. ISSN 1471-2334. doi:

10.1186/s12879-017-2699-8. URL `https://www.ncbi.nlm.nih.gov/pubmed/28893198https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5594572/`.

[423] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jorgensen, W. M. Mooij, B. Muller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Ruger, E. Strand, S. Souissi, R. A. Stillman, R. Vabo, U. Visser, and D. L. DeAngelis. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, 2006. ISSN 0304-3800. doi: 10.1016/j.ecolmodel.2006.04. 023.

[424] World Bank Group. World development indicators, 2020. URL `https://datacatalog.worldbank.org/dataset/world-development-indicators`.

[425] R. H. Wang, Z. Jin, Q. X. Liu, J. van de Koppel, and D. Alonso. A simple stochastic model with environmental transmission explains multi-year periodicity in outbreaks of avian flu. *PLoS One*, 7(2):e28873, 2012. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0028873. URL `https://www.ncbi.nlm.nih.gov/pubmed/22363397`.

[426] C. Valente, J. Hinds, K. A. Gould, F. R. Pinto, H. de Lencastre, and R. Sá-Leão. Impact of the 13-valent pneumococcal conjugate vaccine on Streptococcus pneumoniae multiple serotype carriage. *Vaccine*, 34(34):4072–4078, 2016. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2016.06.017. URL `http://www.sciencedirect.com/science/article/pii/S0264410X16304303`.

[427] J. Mehtälä, M. Antonio, M. S. Kaltoft, K. L. O'Brien, and K. Auranen. Competition between Streptococcus pneumoniae strains: Implications for vaccine-induced replacement in colonization and disease. *Epidemiology*, 24 (4):522–529, 2013. ISSN 1044-3983. doi: 10.1097/EDE.0b013e318294be89. URL `https://journals.lww.com/epidem/Fulltext/2013/07000/Competition_Between_Streptococcus_Pneumoniae.7.aspx`.

[428] H. Adler, D.M. Ferreira, S.B. Gordon, and J. Rylance. Pneumococcal capsular polysaccharide immunity in the elderly. *Clinical and vaccine immunology : CVI*, 24(6):e00004–17, 2017. ISSN 1556-679X 1556-6811. doi: 10.1128/CVI.00004-17. URL https://pubmed.ncbi.nlm.nih.gov/28424198https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5461380/.

[429] O. Le Polain De Waroux, S. Flasche, D. Prieto-Merino, D. Goldblatt, and W. J. Edmunds. The efficacy and duration of protection of pneumococcal conjugate vaccines against nasopharyngeal carriage: A meta-regression model. *Pediatr Infect Dis J*, 34(8):858–64, 2015. ISSN 1532-0987 (Electronic) 0891-3668 (Linking). doi: 10.1097/INF.0000000000000717. URL https://www.ncbi.nlm.nih.gov/pubmed/26075814.

[430] A. Pantosti, D. Boccia, F. D'Ambrosio, S. Recchia, G. Orefici, M. L. Moro, Meningitis National Surveillance of Bacterial, and Study Earss-Italia. Inferring the potential success of pneumococcal vaccination in Italy: serotypes and antibiotic resistance of Streptococcus pneumoniae isolates from invasive diseases. *Microb Drug Resist*, 9 Suppl 1:S61–8, 2003. doi: 10.1089/107662903322541919. URL https://www.ncbi.nlm.nih.gov/pubmed/14633369.

[431] A. B. Suthar, S. D. Lawn, J. del Amo, H. Getahun, C. Dye, D. Sculier, T. R. Sterling, R. E. Chaisson, B. G. Williams, A. D. Harries, and R. M. Granich. Antiretroviral therapy for prevention of tuberculosis in adults with HIV: a systematic review and meta-analysis. *PLoS Med*, 9(7):e1001270, 2012. ISSN 1549-1676 (Electronic) 1549-1277 (Linking). doi: 10.1371/journal.pmed.1001270. URL https://www.ncbi.nlm.nih.gov/pubmed/22911011.

[432] K. J. Rothman, S. Greenland, and T. L. Lash. *Modern epidemiology*. Lippincott Williams & Wilkins, Philadelphia, 3rd edition, 2008. ISBN 9780781755641 0781755646. URL http://www.loc.gov/catdir/enhancements/fy0743/2007036316-d.html.

[433] JJ Deeks, JPT Higgins, and DG Altman. *Chapter 10: Analysing data and undertaking*

meta-analyses, book section 10. Cochrane, 2021. URL `www.training.cochrane.org/handbook`.

[434] S. Balduzzi, G. Rücker, and G. Schwarzer. How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health*, pages 153–160, 2019.

[435] A. Løchen, 2021. URL `https://github.com/alochen/pneumo_ibm`.

[436] A. Melegaro, N. J. Gay, and G. F. Medley. Estimating the transmission parameters of pneumococcal carriage in households. *Epidemiol Infect*, 132(3):433–41, 2004. URL `https://www.ncbi.nlm.nih.gov/pubmed/15188713`.

[437] M. R. Moore, R. Link-Gelles, W. Schaffner, R. Lynfield, C. Holtzman, L. H. Harrison, S. M. Zansky, J. B. Rosen, A. Reingold, K. Scherzinger, A. Thomas, R. E. Guevara, T. Motala, J. Eason, M. Barnes, S. Petit, M. M. Farley, L. McGee, J. H. Jorgensen, and C. G. Whitney. Effectiveness of 13-valent pneumococcal conjugate vaccine for prevention of invasive pneumococcal disease in children in the USA: a matched case-control study. *Lancet Respir Med*, 4(5):399–406, 2016. doi: 10.1016/S2213-2600(16)00052-7. URL `https://www.ncbi.nlm.nih.gov/pubmed/26987984`.

[438] S. Jayasinghe, C. Chiu, H. Quinn, R. Menzies, R. Gilmour, and P. McIntyre. Effectiveness of 7- and 13-valent pneumococcal conjugate vaccines in a schedule without a booster dose: A 10-year observational study. *Clin Infect Dis*, 67(3):367–374, 2018. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1093/cid/ciy129. URL `https://www.ncbi.nlm.nih.gov/pubmed/29471432`.

[439] M. Berman-Rosa, S. O'Donnell, M. Barker, and C. Quach. Efficacy and effectiveness of the PCV-10 and PCV-13 vaccines against invasive pneumococcal disease. *Pediatrics*, 145 (4), 2020. ISSN 1098-4275 (Electronic) 0031-4005 (Linking). doi: 10.1542/peds.2019-0377. URL `https://www.ncbi.nlm.nih.gov/pubmed/32156773`.

[440] S. A. Madhi and M. D. Knoll. An affordable pneumococcal conjugate vaccine after 20 years. *Lancet Infect Dis*, 21(6):751–753, 2021. ISSN 1474-4457 (Electronic) 1473-3099

(Linking). doi: 10.1016/S1473-3099(21)00002-5. URL `https://www.ncbi.nlm.nih.gov/pubmed/33516294`.

[441] W. P. Hanage, J. A. Finkelstein, S. S. Huang, S. I. Pelton, A. E. Stevenson, K. Kleinman, V. L. Hinrichsen, and C. Fraser. Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete. *Epidemics*, 2(2):80–4, 2010. ISSN 1878-0067 (Electronic) 1878-0067 (Linking). doi: 10.1016/j.epidem.2010.03.005. URL `https://www.ncbi.nlm.nih.gov/pubmed/21031138`.

[442] S. F. Dowell, C. G. Whitney, C. Wright, Jr. Rose, C. E., and A. Schuchat. Seasonal patterns of invasive pneumococcal disease. *Emerg Infect Dis*, 9(5):573–9, 2003. doi: 10.3201/eid0905.020556. URL `https://www.ncbi.nlm.nih.gov/pubmed/12737741`.

[443] A. Morimura, S. Hamaguchi, Y. Akeda, and K. Tomono. Mechanisms underlying pneumococcal transmission and factors influencing host-pneumococcus interaction: A review. *Front Cell Infect Microbiol*, 11:639450, 2021. ISSN 2235-2988 (Electronic) 2235-2988 (Linking). doi: 10.3389/fcimb.2021.639450. URL `https://www.ncbi.nlm.nih.gov/pubmed/33996623`.

[444] C. Valente, J. Hinds, F. Pinto, S. D. Brugger, K. Gould, K. Muhlemann, H. de Lencastre, and R. Sa-Leao. Decrease in pneumococcal co-colonization following vaccination with the seven-valent pneumococcal conjugate vaccine. *PLoS One*, 7(1):e30235, 2012. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0030235. URL `https://www.ncbi.nlm.nih.gov/pubmed/22253924`.

[445] J. Chan, C. D. Nguyen, E. M. Dunne, E. Kim Mulholland, T. Mungun, W. S. Pomat, E. Rafai, C. Satzke, D. M. Weinberger, and F. M. Russell. Using pneumococcal carriage studies to monitor vaccine impact in low- and middle-income countries. *Vaccine*, 37 (43):6299–6309, 2019. doi: 10.1016/j.vaccine.2019.08.073. URL `https://www.ncbi.nlm.nih.gov/pubmed/31500968`.

[446] S. N. Ladhani, N. J. Andrews, P. Waight, R. Borrow, M. P. Slack, and E. Miller. Invasive pneumococcal disease, comorbidities, and polysaccharide vaccine use in children

aged 5-15 years in England and Wales. *Clin Infect Dis*, 58(4):517–25, 2014. ISSN 1537-6591 (Electronic) 1058-4838 (Linking). doi: 10.1093/cid/cit791. URL `https://www.ncbi.nlm.nih.gov/pubmed/24300040`.

[447] A. de St Maurice, W. Schaffner, M. R. Griffin, N. Halasa, and C. G. Grijalva. Persistent sex disparities in invasive pneumococcal diseases in the conjugate vaccine era. *J Infect Dis*, 214(5):792–7, 2016. ISSN 1537-6613 (Electronic) 0022-1899 (Linking). doi: 10.1093/infdis/jiw222. URL `https://www.ncbi.nlm.nih.gov/pubmed/27247342`.

[448] G. H. Wagenvoort, E. A. Sanders, B. J. Vlaminckx, H. E. de Melker, A. van der Ende, and M. J. Knol. Sex differences in invasive pneumococcal disease and the impact of pneumococcal conjugate vaccination in the Netherlands, 2004 to 2015. *Euro Surveill*, 22 (10), 2017. ISSN 1560-7917 (Electronic) 1025-496X (Linking). doi: 10.2807/1560-7917.ES. 2017.22.10.30481. URL `https://www.ncbi.nlm.nih.gov/pubmed/28300529`.

[449] K. L. O'Brien, H. C. Baggett, W. A. Brooks, D. R. Feikin, L. L. Hammitt, M. M. Higdon, S. R. C. Howie, M. D. Knoll, K. L. Kotloff, O. S. Levine, S. A. Madhi, D. R. Murdoch, C. Prosperi, J. A. G. Scott, Q. Shi, D. M. Thea, Z. Wu, S. L. Zeger, P. V. Adrian, P. Akarasewi, T. P. Anderson, M. Antonio, J. O. Awori, V. L. Baillie, C. Bunthi, J. Chipeta, M. J. Chisti, J. Crawley, A. N. DeLuca, A. J. Driscoll, B. E. Ebruke, H. P. Endtz, N. Fancourt, W. Fu, D. Goswami, M. J. Groome, M. Haddix, L. Hossain, Y. Jahan, E. W. Kagucia, A. Kamau, R. A. Karron, S. Kazungu, N. Kourouma, L. Kuwanda, G. Kwenda, M. Li, E. M. Machuka, G. Mackenzie, N. Mahomed, S. A. Maloney, J. L. McLellan, J. L. Mitchell, D. P. Moore, S. C. Morpeth, A. Mudau, L. Mwananyanda, J. Mwansa, M. Silaba Ominde, U. Onwuchekwa, D. E. Park, J. Rhodes, P. Sawatwong, P. Seidenberg, A. Shamsul, E. A. F. Simões, S. Sissoko, S. Wa Somwe, S. O. Sow, M. Sylla, B. Tamboura, M. D. Tapia, S. Thamthitiwat, A. Toure, N. L. Watson, K. Zaman, and S. M. A. Zaman. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *The Lancet*, 394(10200):757–779, 2019. ISSN 0140-6736. doi: 10.1016/S0140-6736(19)30721-4. URL `https://doi.org/10.1016/S0140-6736(19)30721-4`.

[450] European Centre for Disease Prevention and Control. Invasive pneumococcal disease, 2019. URL `https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-invasive-pneumococcal-disease.pdf`.

[451] M. Deloria Knoll, J. C. Bennett, M. Garcia Quesada, E. W. Kagucia, M. E. Peterson, D. R. Feikin, A. L. Cohen, M. K. Hetrich, Y. Yang, J. N. Sinkevitch, K. Ampofo, L. Aukes, S. Bacci, G. Bigogo, M. C. Brandileone, M. G. Bruce, R. Camilli, J. Castilla, G. Chan, G. Chanto Chacón, P. Ciruela, H. Cook, M. Corcoran, R. Dagan, K. Danis, S. de Miguel, P. De Wals, S. Desmet, Y. Galloway, T. Georgakopoulou, L. L. Hammitt, M. Hilty, P. Ho, S. Jayasinghe, J. D. Kellner, J. Kleynhans, M. J. Knol, J. Kozakova, K. G. Kristinsson, S. N. Ladhani, C. S. Lara, M. E. León, T. Lepp, G. A. Mackenzie, L. Mad'arová, A. McGeer, T. Mungun, J. M. Mwenda, J. P. Nuorti, N. Nzoyikorera, K. Oishi, L. H. De Oliveira, M. Paragi, T. Pilishvili, R. Puentes, E. Rafai, S. K. Saha, L. Savrasova, C. Savulescu, J. A. Scott, K. J. Scott, F. Serhan, L. P. Setchanova, N. Sinkovec Zorko, A. Skoczyńska, T. D. Swarthout, P. Valentiner-Branth, M. van der Linden, D. F. Vestrheim, A. von Gottberg, I. Yildirim, K. Hayford, and The Pserenade Team. Global landscape review of serotype-specific invasive pneumococcal disease surveillance among countries using PCV10/13: The pneumococcal serotype replacement and distribution estimation (PSERENADE) project. *Microorganisms*, 9(4):742, 2021. ISSN 2076-2607. doi: 10.3390/microorganisms9040742. URL `https://pubmed.ncbi.nlm.nih.gov/33918127https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8066045/`.

[452] S. C. Inzaule, S. K. Tessema, Y. Kebede, A. E. Ogwell Ouma, and J. N. Nkengasong. Genomic-informed pathogen surveillance in Africa: opportunities and challenges. *Lancet Infect Dis*, 2021. ISSN 1474-4457 (Electronic) 1473-3099 (Linking). doi: 10.1016/S1473-3099(20)30939-7. URL `https://www.ncbi.nlm.nih.gov/pubmed/33587898`.

[453] Sampling map, 2016. URL `https://www.pneumogen.net/gps/sampling_map.html`.

[454] J. Reyes-Silveyra and A. R. Mikler. Modeling immune response and its effect on

infectious disease outbreak dynamics. *Theoretical Biology and Medical Modelling*, 13(1): 10, 2016. ISSN 1742-4682. doi: 10.1186/s12976-016-0033-6. URL `https://doi.org/10.1186/s12976-016-0033-6`.

[455] C. R. MacIntyre and J. Leask. Immunization myths and realities: responding to arguments against immunization. *J Paediatr Child Health*, 39(7):487–91, 2003. doi: 10.1046/j.1440-1754.2003.t01-1-00200.x. URL `https://www.ncbi.nlm.nih.gov/pubmed/12969200`.

[456] S. E. Londono, X. Li, J. Toor, M. J. de Villiers, S. Nayagam, T. B. Hallett, K. Abbas, M. Jit, P. Klepac, K. Jean, T. Garske, N. M. Ferguson, and K. A. M. Gaythorpe. How can the public health impact of vaccination be estimated? *medRxiv*, page 2021.01.08.21249378, 2021. doi: 10.1101/2021.01.08.21249378. URL `https://www.medrxiv.org/content/medrxiv/early/2021/01/09/2021.01.08.21249378.full.pdf`.

[457] J. N. Weiser, D. M. Ferreira, and J. C. Paton. Streptococcus pneumoniae: transmission, colonization and invasion. *Nat Rev Microbiol*, 16(6):355–367, 2018. ISSN 1740-1534 (Electronic) 1740-1526 (Linking). doi: 10.1038/s41579-018-0001-8. URL `https://www.ncbi.nlm.nih.gov/pubmed/29599457`.

[458] World Health Organization. Global market study: Pneumococcal conjugate (PCV) and polysaccharide (PPV) vaccines working document, 2020. URL `https://www.who.int/immunization/programmes_systems/procurement/mi4a/platform/module2/Pneumococcal_Vaccine_Market_Study-June2020.pdf?ua=1`.

[459] H. E. Hsu, K. A. Shutt, M. R. Moore, B. W. Beall, N. M. Bennett, A. S. Craig, M. M. Farley, J. H. Jorgensen, C. A. Lexau, S. Petit, A. Reingold, W. Schaffner, A. Thomas, C. G. Whitney, and L. H. Harrison. Effect of pneumococcal conjugate vaccine on pneumococcal meningitis. *N Engl J Med*, 360(3):244–56, 2009. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/NEJMoa0800836. URL `https://www.ncbi.nlm.nih.gov/pubmed/19144940`.

[460] K. P. Klugman and S. Black. Impact of existing vaccines in reducing antibiotic resistance:

Primary and secondary effects. *Proc Natl Acad Sci U S A*, 115(51):12896–12901, 2018. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1721095115. URL `https://www.ncbi.nlm.nih.gov/pubmed/30559195`.

[461] A. J. van Hoek, N. Andrews, P. A. Waight, R. George, and E. Miller. Effect of serotype on focus and mortality of invasive pneumococcal disease: Coverage of different vaccines and insight into non-vaccine serotypes. *PLOS ONE*, 7(7):e39150, 2012. doi: 10.1371/journal.pone.0039150. URL `https://doi.org/10.1371/journal.pone.0039150`.

[462] J.Y. Baek, S.H. Kim, CI. Kang, DR. Chung, K.R. Peck, K.S. Ko, and JH. Song. Prevalence of antimicrobial resistant Streptococcus pneumoniae serotype 11A isolates in Korea, during 2004–2013, due to the increase of multidrug-resistant clone, CC166. *Infection, Genetics and Evolution*, 38:122–125, 2016. ISSN 1567-1348. doi: https://doi.org/10.1016/j.meegid.2015.12.018. URL `https://www.sciencedirect.com/science/article/pii/S1567134815300836`.

[463] B. Hayman and S. Pagliusi. Emerging vaccine manufacturers are innovating for the next decade. *Vaccine X*, 5:100066, 2020. ISSN 2590-1362 (Electronic) 2590-1362 (Linking). doi: 10.1016/j.jvacx.2020.100066. URL `https://www.ncbi.nlm.nih.gov/pubmed/32462140`.

[464] M. R. Alderson, V. Sethna, L. C. Newhouse, S. Lamola, and R. Dhere. Development strategy and lessons learned for a 10-valent pneumococcal conjugate vaccine (Pneumosil). *Hum Vaccin Immunother*, pages 1–8, 2021. ISSN 2164-554X (Electronic) 2164-5515 (Linking). doi: 10.1080/21645515.2021.1874219. URL `https://www.ncbi.nlm.nih.gov/pubmed/33625961`.

[465] M. Reglinski, G. Ercoli, C. Plumptre, E. Kay, F. C. Petersen, J. C. Paton, B. W. Wren, and J. S. Brown. A recombinant conjugated pneumococcal vaccine that protects against murine infections with a similar efficacy to Prevnar-13. *NPJ Vaccines*, 3:53, 2018. ISSN

2059-0105 (Electronic) 2059-0105 (Linking). doi: 10.1038/s41541-018-0090-4. URL `https://www.ncbi.nlm.nih.gov/pubmed/30393571`.

[466] E. Kay, J. Cuccui, and B. W. Wren. Recent advances in the production of recombinant glycoconjugate vaccines. *NPJ Vaccines*, 4:16, 2019. ISSN 2059-0105 (Electronic) 2059-0105 (Linking). doi: 10.1038/s41541-019-0110-z. URL `https://www.ncbi.nlm.nih.gov/pubmed/31069118`.

[467] E. J. Kay, L. E. Yates, V. S. Terra, J. Cuccui, and B. W. Wren. Recombinant expression of Streptococcus pneumoniae capsular polysaccharides in Escherichia coli. *Open Biol*, 6(4): 150243, 2016. ISSN 2046-2441 (Electronic) 2046-2441 (Linking). doi: 10.1098/rsob.150243. URL `https://www.ncbi.nlm.nih.gov/pubmed/27110302`.

[468] J. J. Chen, L. Yuan, Z. Huang, N. M. Shi, Y. L. Zhao, S. L. Xia, G. H. Li, R. C. Li, Y. P. Li, S. Y. Yang, and J. L. Xia. Safety and immunogenicity of a new 13-valent pneumococcal conjugate vaccine versus a licensed 7-valent pneumococcal conjugate vaccine: a study protocol of a randomised non-inferiority trial in China. *BMJ open*, 6(10):e012488–e012488, 2016. ISSN 2044-6055. doi: 10.1136/bmjopen-2016-012488. URL `https://pubmed.ncbi.nlm.nih.gov/27798013https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5073664/`.

[469] Walvax Biotechnology. Immunogenicity and safety study of the 13-valent pneumococcal polysaccharide conjugate vaccine in 2-71 months old healthy infants and toddlers (the youngest could be 6 weeks old), 2020. URL `https://ClinicalTrials.gov/show/NCT02736240`.

[470] Pfizer. Trial to evaluate the safety and immunogenicity of a 20-valent pneumococcal conjugate vaccine in pneumococcal vaccine-naïve adults, 2020. URL `https://ClinicalTrials.gov/show/NCT03760146`.

[471] Merck. A study to evaluate the safety, tolerability, and immunogenicity of V114 followed by PNEUMOVAX™23 in adults at increased risk for pneumococcal dis-

ease (V114-017/PNEU-DAY), 2021. URL `https://ClinicalTrials.gov/show/NCT03547167`.

[472] Merck. A study to evaluate the safety, tolerability and immunogenicity of V114 in healthy adults and infants (V114-005), 2021. URL `https://ClinicalTrials.gov/show/NCT02531373`.

[473] Beijing Minhai Biotechnology Co. A phase III clinical trial of a 13-valent pneumococcal conjugate vaccine in healthy infants, 2019. URL `https://ClinicalTrials.gov/show/NCT02494999`.

[474] GlaxoSmithKline. Safety and immunogenicity study of GlaxoSmithKline (GSK) biologicals' pneumococcal vaccine 2830930A administered in toddlers, 2014. URL `https://ClinicalTrials.gov/show/NCT01485406`.

[475] M. Horn, U. Behre, M. Traskine, K. Dobbelaere, and D. Borys. Safety, reactogenicity, and immunogenicity of a 12-valent pneumococcal non-typeable Haemophilus influenzae protein D-conjugate vaccine in healthy toddlers: results from a phase I, randomized trial. *Hum Vaccin Immunother*, 17(5):1463–1469, 2021. ISSN 2164-554X (Electronic) 2164-5515 (Linking). doi: 10.1080/21645515.2020.1810493. URL `https://www.ncbi.nlm.nih.gov/pubmed/33175600`.

[476] M. R. Alderson. Status of research and development of pediatric vaccines for Streptococcus pneumoniae. *Vaccine*, 34(26):2959–2961, 2016. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2016.03.107. URL `https://www.ncbi.nlm.nih.gov/pubmed/27083428`.

[477] EuBiologics Co. Safety and immunogenicity of the 'EuPCV15' in healthy Korean adults, 2021. URL `https://ClinicalTrials.gov/show/NCT04830358`.

[478] C. Lee, S. K. Choi, R. K. Kim, H. Kim, Y. H. Whang, H. Pharm, H. Cheon, D. Y. Yoon, C. W. Kim, Y. O. Baik, S. S. Park, and I. Lee. Development of a new 15-valent pneumococcal conjugate vaccine (PCV15) and evaluation of its immunogenicity. *Biologicals*, 61:

32–37, 2019. ISSN 1095-8320 (Electronic) 1045-1056 (Linking). doi: 10.1016/j.biologicals. 2019.07.005. URL `https://www.ncbi.nlm.nih.gov/pubmed/31416790`.

[479] A. E. B. Turner, J. E. Gerson, H. Y. So, D. J. Krasznai, A. J. St Hilaire, and D. F. Gerson. Novel polysaccharide-protein conjugates provide an immunogenic 13-valent pneumococcal conjugate vaccine for S. pneumoniae. *Synth Syst Biotechnol*, 2(1):49–58, 2017. doi: 10.1016/j.synbio.2016.12.002. URL `https://www.ncbi.nlm.nih.gov/pubmed/29062961`.

[480] S. Flasche. The scope for pneumococcal vaccines that do not prevent transmission. *Vaccine*, 35(45):6043–6046, 2017. ISSN 1873-2518 (Electronic) 0264-410X (Linking). doi: 10.1016/j.vaccine.2017.09.073. URL `https://www.ncbi.nlm.nih.gov/pubmed/28982625`.

[481] U. Obolski, J. Lourenco, C. Thompson, R. Thompson, A. Gori, and S. Gupta. Vaccination can drive an increase in frequencies of antibiotic resistance among nonvaccine serotypes of Streptococcus pneumoniae. *Proc Natl Acad Sci U S A*, 115(12):3102–3107, 2018. doi: 10.1073/pnas.1718712115. URL `https://www.ncbi.nlm.nih.gov/pubmed/29511100`.

[482] R. J. Mostowy, N. J. Croucher, N. De Maio, C. Chewapreecha, S. J. Salter, P. Turner, D. M. Aanensen, S. D. Bentley, X. Didelot, and C. Fraser. Pneumococcal capsule synthesis locus cps as evolutionary hotspot with potential to generate novel serotypes by recombination. *Mol Biol Evol*, 34(10):2537–2554, 2017. ISSN 1537-1719 (Electronic) 0737-4038 (Linking). doi: 10.1093/molbev/msx173. URL `https://www.ncbi.nlm.nih.gov/pubmed/28595308`.

Appendix A

Supplementary material for systematic review of dynamic transmission models and economic evaluations

ISPOR checklist

Relevance		
1	Is the population relevant?	Are the demographics similar?
		Are risk factors similar?
		Are behaviors similar?
		Is the medical condition similar?
		Are comorbidities similar?
2	Are any critical interventions missing?	Does the intervention analyzed in the model match the intervention you are interested in?
		Have all relevant comparators been considered?
		Does the background care in the model match yours?
3	Are any relevant outcomes missing?	Are the health outcomes relevant to you considered?
		Are the economic end points relevant to you considered?
4	Is the context (settings and circumstances) applicable?	Is the geographic location similar?
		Is the health care system similar?
		Is the time horizon applicable to your decision?
		Is the analytic perspective appropriate to your decision problem?
Credibility		
Validation		
1	Is external validation of the model sufficient to make its results credible for your decision?	Has the model been shown to accurately reproduce what was observed in the data used to create the model?
		Has the model been shown to accurately estimate what actually happened in one or more separate studies?
		Has the model been shown to accurately forecast what eventually happens in reality?
2	Is internal verification of the model sufficient to make its results credible for your decision?	Have the process of internal verification and its results been documented in detail?
		Has the testing been performed systematically?
		Does the testing indicate that all the equations are consistent with their data sources?
		Does the testing indicate that the coding has been correctly implemented?
3	Does the model have sufficient face validity to make its results credible for your decision?	Does the model contain all the aspects considered relevant to the decision?

		Are all the relevant aspects represented and linked according to the best understanding of their characteristics?
		Have the best available data sources been used to inform the various aspects?
		Is the time horizon sufficiently long to account for all relevant aspects of the decision problem?
		Are the results plausible?
		If others have rated the face validity, did they have a stake in the results?
Design		
4	Is the design of the model adequate for your decision problem?	Was there a clear, written statement of the decision problem, modeling objective, and scope of the model?
		Was there a formal process for developing the model design (e.g. influence diagram, concept map)?
		Is the model concept and structure consistent with, and adequate to address, the decision problem/objective and the policy context?
		Have any assumptions implied by the design of the model been described, and are they reasonable for your decision problem?
		Is the choice of model type appropriate?
		Were key uncertainties in model structure identified and their implications discussed?
Data		
5	Are the data used in populating the model suitable for your decision problem?	All things considered, do you agree with the values used for the inputs?
		Did the approaches to obtaining and processing the data inputs meet the criteria from their corresponding questionnaires?
Analysis		
6	Were the analyses performed using the model adequate to inform your decision problem?	
7	Was there an adequate assessment of the effects of uncertainty?	
Reporting		
8	Was the reporting of the model adequate to inform your decision problem?	Did the report of the analyses provide the results needed for your decision problem?
		Was adequate nontechnical documentation freely accessible to any interested reader?
		Was technical documentation, in sufficient detail to allow (potentially) for replication, made

		available openly or under agreements that protect intellectual property?
Interpretation		
9	Was the interpretation of results fair and balanced?	
Conflict of Interest		
10	Were there any potential conflicts of interest?	
11	If there were potential conflicts of interest, were steps taken to address these?	

Table A.1: ISPOR results for included papers

Economic Evaluations		Ansaldi 2020	Bakir 2012	Beutels 2011	Blank 2015	By 2012	Castiglia 2017	Chuck 2010	De Wals 2009	Delgeize 2016	Earnshaw 2012	Klok 2013	Knerer 2012	Kuhlmann 2017	Luca 2018	Nakamura 2011	Pugh 2020	Robberstad 2011	Rozenbaum 2010	Rubin 2010	Smith 2021	Strutton 2012	Talbird 2010	Vemer 2014	Vucina 2015	Waye 2015	Wilson 2018
Relevance	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	3	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	4	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Assessment	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
Credibility — Validation	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2	Y	Y[1]	Y	N	Y	Y[1]	Y[1]	Y	Y[1]	N	N[1]	Y	Y	Y	N[2]	Y	Y	Y[1]	Y	Y	N	Y[1]	Y[1]	Y[1]	N	Y
	3	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Assessment	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne
Credibility — Design	4	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	Assessment	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne
Data	5	N	N	Y	N	Y	Y	N	N	Y	N	N	N	Y	Y	N	N	N	N	Y	Y	N	N	Y	Y	N	Y
	Assessment	W	W	Ne	W	Ne	Ne	W	W	Ne	W	W	W	Ne	Ne	W	W	W	W	Ne	Ne	W	W	Ne	Ne	W	Ne

350

Economic Evaluations

Credibility		Ansaldi 2020	Bakir 2012	Beutels 2011	Blank 2012	By 2012	Castiglia 2017	Chuck 2010	De Wals 2009	Delgeize 2016	Earnshaw 2012	Klok 2013	Knerer 2012	Kuhlmann 2017	Luca 2018	Nakamura 2011	Pugh 2020	Robberstad 2011	Rozenbaum 2010	Rubin 2010	Smith 2021	Strutton 2012	Talbird 2010	Vemer 2014	Vucina 2015	Waye 2015	Wilson 2018
Analysis	6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Analysis	7	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Analysis	Assessment	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Ne	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str
Interpret Reporting	8	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Interpret Reporting	Assessment	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str	Str
COI	9	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
COI	Assessment	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne	Ne
COI	10	Y	Y	Y	Y	Y	Y	NA[3]	NA[3]	Y	Y	Y	Y	NA[3]	NA[3]	NA[3]	NA[3]	Y	Y	Y	Y	Y	Y	Y	N	NA[3]	NA[3]
COI	11	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N
COI	Assessment	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	Str	W	W	W	W	W	W
	Assessment	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S

N: No; Ne: Neutral; S: Sufficient; **Str**: Strength; **W**: Weakness; **Y**: Yes.

[1] based on De Wals 2009 (Bakir 2012, Talbird 2010), Knerer 2012 (Castiglia 2017, Delgeize 2016), Bedard 2009 (Chuck 2010, Strutton 2012 (Klok 2013), Bos 2003 (Rozenbaum 2010, Vemer 2014), TRIVAC/Clark 2013 (Vucina 2015)

[2] Experts provided advice but details on this process not provided

[3] Not reported

Table A.2: Summary of costs and perspectives in economic evaluations of PCV. When hospitalization costs listed both children and adults, the higher of the two is listed here.

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Ansaldi 2020	EUR	Healthcare/payer	€5	€20 difference between PCV13 and PCV10	3% for both costs and health outcomes	BAC: €3176 - €5493; MEN: €51 - €2190; PNE: €34 - €1079; AOM: €76
Bakir 2012	N/A USD	Healthcare	$3.26/dose	$30/dose[2]	N/A	MEN: $1000; BAC: $500; PNE: $325; AOM: $438
Beutels 2011	2010 EUR	Healthcare	€5/dose	PCV7: €66.15/dose PCV10: €70.44/dose PCV13: €74.55/dose	Costs 3% and health outcomes 1.5%	MEN: €6306; BAC: €6282; Other IPD: €3108; PNE: €4775; AOM: €2946
Blank 2012	2012 EUR	Government	€24/dose	PCV7: €58/dose PCV13: €74/dose	3% for both costs and health outcomes	MEN: €3256; BAC: €1378; PNE: €2820; AOM: €426
By 2012	2010 SEK	Societal	18 SEK/dose	518.95 SEK/dose[2]	3% for both costs and health outcomes	MEN: 151,348 SEK; BAC: 73,577 SEK; PNE: 31,866 SEK; AOM: 9000 SEK
Castiglia 2017	2015 EUR	Healthcare	N/A[3]	€47.73/dose[2]	3% for both costs and health outcomes	MEN: €8067; BAC: €3176; PNE: €1948; AOM: €662

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Chuck 2010	2008 CAN$	Healthcare	N/A[3]	$284 (3+1)	Undiscounted	MEN: CAN$35,017; BAC: CAN$11,695; PNE: CAN$8031; AOM: CAN$616
De Wals 2009	2001 CAN$	Societal, Healthcare, Family	N/A[3]	$71/dose[2]	3% for both costs and health outcomes	IPD: CAN$; PNE: CAN$; AOM: CAN$
Delgleize 2016	2014 GBP	Healthcare	£7.64/dose	£48.88/dose[2]	3.5% for both costs and health outcomes	MEN: £7872; BAC: £6595; PNE: £5226; AOM: £1142
Earnshaw 2012	2010 CAN$	Healthcare	$7.84/dose	$86.26/dose[2]	N/A	MEN: CAN$61,115; BAC: CAN$13,560; PNE: CAN$10,305; AOM: CAN$95
Klok 2013	2011 DKK; 2010 SEK	Payer	15.65 DKK 18 SEK	451.33 DKK[2] 518.95 SEK[2]	3% for health outcomes only but not costs	MEN: 61,702 DKK; BAC: 41,613 DKK; PNE: 23,582 DKK; AOM: 7741 DKK; MEN: 71,888 SEK; BAC: 67,454 SEK; PNE: 32,562 SEK; AOM: 8345 SEK

353

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Knerer 2012	2007 GBP, CAN$	Societal, Healthcare, Family	N/A[3]	£27.60/dose (UK)[2] $70/dose (Canada)[2]	3.5% (UK) and 3% (Canada) for both costs and health outcomes	UK MEN: £7504; BAC: £5360; PNE: £4457; AOM: £722; Canada MEN: CAN$31,792; BAC: CAN$11,034; PNE: CAN$7305; AOM: CAN$585
Kuhlmann 2017	2013 EUR	Societal, Healthcare	€6.95/dose	PCV10: €52.57/dose PCV13: €55.82/dose	3% for both costs and health outcomes	IPD: €9906; PNE: €8372; AOM: €1392
Luca 2018	2013 CAN$	Healthcare	N/A	N/A	N/A	PNE: $205.8 million (total); Other: $1648.1 million (total)
Nakamura 2011	2005 USD	Societal	$5/dose	$10/dose for lower middle-income countries; $20/dose for higher middle-income countries1	4% for both costs and health outcomes	MEN: $5350-$5657; NPNM: $1105-$1112; PNE: $1301-$2213

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Pugh 2020	2016/2017 EUR	Payer	Finland: €5.00/dose; NL: €10.63/dose	PCV10: €44.00/dose (Finland), €57.13/dose (NL); PCV13 €58.41/dose (Finland), €68.56/dose (NL)	Finland: 3% for both costs and health outcomes; NL: 4.0% for costs and 1.5% for health outcomes	**Finland:** BAC: €2065 - €7095/case; MEN: €22,949-€22,387/case; PNE: €187-€6000/case; AOM: €103/case; **Netherlands:** BAC: €5572-€5928/case; MEN: €10,162 - €17,653/case; PNE: €514-€6112/case; AOM: €20/case
Robberstad 2011	2009 NOK	Societal	81 NOK/dose	346 NOK/dose[2]	4% for both costs and health outcomes	MEN: 117,156 NOK; BAC: 106,488 NOK; PNE: 45,270 NOK; AOM: 9782 NOK
Rozenbaum 2010	EUR	Healthcare	€5.95/dose	PCV7: €50/dose PCV10: €62.25/dose PCV13: €68.56/dose	Costs 4% and health effects 1.5%	IPD: €1091 - 27,318; PNE: €26 - 2614; AOM: €17 - 381
Rubin 2010	2008 USD	Societal	$11/dose	PCV7: $73/dose PCV13: $100/dose	3% for both costs and health outcomes	MEN: $17,048; BAC: $3253; PNE: $10,148; AOM: $840

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Smith 2021	2017 USD	Healthcare	$25.84 (21.33-31.18)	PCV13: $188; PCV15: $199; PCV20: $222; Hypothetical adult PCV20: $255	3% for both costs and health outcomes	IPD: $21,425-$43,628; PNE: $18,050-$35,950; Disability: $14,239/year
Strutton 2012	2008/ 2009 EUR	Healthcare	€7/dose (Germany); €30/dose (Greece); €5.95/dose (Netherlands)	PCV7: €49/dose (Germany) €52.40/dose (Greece) €57.13/dose (NL) PCV10: €39.90/dose (Germany) €42.14/dose (Greece) €57.13/dose (NL) PCV13: €49/dose (Germany) €49.64/dose (Greece) €68.56/dose (NL)	5% (Germany), 3% (Greece) and 1.5% (NL) respectively for health outcomes only but not costs	**Germany:** MEN: €10,071; BAC: €8503; PNE: €5627; AOM: €135; **Greece:** MEN: €34,163; BAC: €18,635; PNE: €5708; AOM: €3861; **Netherlands:** MEN: €18,781; BAC: €10,759; PNE: €2615; AOM: €96

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Talbird 2010	2008 CAN$/EUR/Pesos/NOK	Healthcare	N/A	Not considered because of price parity and analysis focused on incremental costs	Undiscounted	**Canada:** MEN: CAN$7213 BAC: CAN$11,048 PNE: CAN$7616 AOM: CAN$593 **Germany:** MEN: €2940; BAC: €3265; PNE: €2789; AOM: €40; **Mexico:** MEN: 172,849 Pesos; BAC: 73,934 Pesos; PNE: 76,217 Pesos; AOM: 5315 Pesos; **Norway:** MEN: 114,733 NOK; BAC: 104,285 NOK; PNE: 44,333 NOK; AOM: 9778 NOK
Vemer 2014	2012 EUR	Societal	N/A[3]	Price difference between vaccines: € 0 -€25/dose (baseline €11)	Costs 4% and health outcomes 1.5%	PNE: €3147.82; AOM: €657.59;
Vucina 2015	2014 USD	Government, Societal	N/A[3]	PCV10: $63/dose; $30/dose[4] PCV13: $67/dose; $35/dose[4]	3% for both costs and health outcomes	MEN: $8000; NPNM: $4000; PNE: $2000; Household: $50 (all)

Authors and year published	Price year	Perspective	Administrative cost	Vaccine cost	Discounting rate	Hospitalization cost per case[1]
Waye 2015	2014 CAN$	Direct payer	N/A	Not considered because of price parity	3% for both costs and health outcomes	MEN: CAN$ 38,070; BAC: CAN$ 12,714; PNE: CAN$ 8731
Wilson 2018	2017 CAN$	Societal	CAN$ 6.63	PCV10 estimated to be 30% lower than PCV13	3% for both costs and health outcomes	BAC: CAN$18,820-32,274; MEN: CAN$23,434-42,911; PNE: CAN$118.55-10,699

[1] AOM: acute otitis media. BAC: bacteraemia. CAP: community-acquired pneumonia. LY: life years. MEN: meningitis. NPNM: Pneumococcal non-pneumonia non-meningitis. NTHi: non-typeable Haemophilus influenzae. PNE: pneumonia. QALY: quality-adjusted life years.

[2] Assumed price parity (PCV7 = PCV10 = PCV13)

[3] It was assumed that if no administration cost was provided, it was accounted for in the vaccine cost listed

[4] With universal childhood immunization

358

Appendix B

ODD protocol for individual-based dynamic transmission model

IBM ODD standard protocol

1. Purpose

The purpose of the model was to evaluate the epidemiological impact of increasing serotype valency in the pneumococcal conjugate vaccines in two countries: France and the USA. The model predicts the number of carriers and disease cases before, during and after vaccination with expanded PCVs over 50 years. For this, longitudinal carriage data from naïve children in a Thai refugee camp was used. Additionally, previously obtained invasiveness estimates in those locations were used to convert the number of carriers (source of transmission) to disease cases (vaccine clinical endpoints). An IBM was used because of the complexity of age-stratified compartmental models with increasing number of serotypes.

2. State variables and scales

The model describes 5000 individuals who are either susceptible (0) to or carrying (1) one of *n* serotypes. Each individual is characterized by their age (in days) which classifies them as children, adults or elderly, and can, depending on whether they are a carrier or not, either recover from carriage (1 → 0) at the recovery rate or get infected with the serotype (0 →1) at the force of infection (FOI). FOI is defined as the product of the transmission (or contact) rate β, the number of infected people and the number of contacts per time between age groups.

2.1 Entities

Individual human

Variable	States
Age	Continuous
Epidemiological state for each serotype	0 (uninfected but susceptible) or 1 (carrier)
Number of serotypes carried	0 – 2 (max carrying capacity is 2 serotypes)

Serotype

| Variable	Values
Recovery rate (= 1/duration of carriage)	Serotype-specific; fitted to Thai refugee camp data
FOI	Fitted to obtain 40% carriage prevalence in original naïve population; 0.000121
Number of carriers	Varies over time
Invasiveness	Serotype-specific invasiveness from either local (country-specific) invasiveness dataset or global (fitted across multiple datasets) invasiveness dataset

Vaccine

Variable	Values
Duration of protection	15 years
Efficacy	95%
Number of serotypes within vaccine	13, 20, 30

3. Process overview and scheduling

The processes in the model occur as part of a continuous time system, where people age and are born/die, and a switching event takes place at each time step. The time step is fixed at 1 day and a fixed number of events happen each time step. This was fixed to about (# of individuals)/50 events per time step based on trial and error, such that enough events occurred to propagate infection in a naïve child population without serotypes becoming extinct. The simulation runs from 20 years pre-vaccination to 10 years post-vaccination with each vaccine. The model works as follows: first, for each serotype, the total FOI was calculated for the whole population. For each serotype, each individual was assigned either the FOI rate if they were susceptible or the recovery rate if they were carriers. To choose the person that the switching

event happens to, a random number is drawn from a uniform distribution $x \in U(0,1)$ and multiplied by the sum of all individual rates. When compared to the cumulative sum of individual rates, this number will be in the interval of the rate corresponding to the chosen person. The chosen person will then have their serotype rates changed depending on how many serotypes they are carrying at the time they were chosen. If they are carrying 2 serotypes (i.e. the maximum carrying capacity), then all their other serotype FOI rates become 0 such that no other serotypes can be acquired. The maximum carrying capacity was chosen to be 2 to avoid complexity and because this is the most common number of serotypes detected in multi-serotype carriage [25]. The switching event happens to the serotype chosen using the same method as the person-choosing. There is evidence for two mechanisms of immunity to pneumococcus: anticapsular (serotype-specific) immunity and non-specific [26-29]. Serotype-specific immunity reduces the likelihood of re-acquisition of a certain serotype, enhancing serotype competition and potentially providing a mechanism to maintain serotype diversity. In this model, it is a reduction in the FOI for the following colonizing serotypes of that type. Non-specific immunity is evidenced by the higher disease incidence in children, particularly infants. In this model, it is a reduction in the duration of carriage. The Unified Modelling Language (UML) diagram of this process is below.
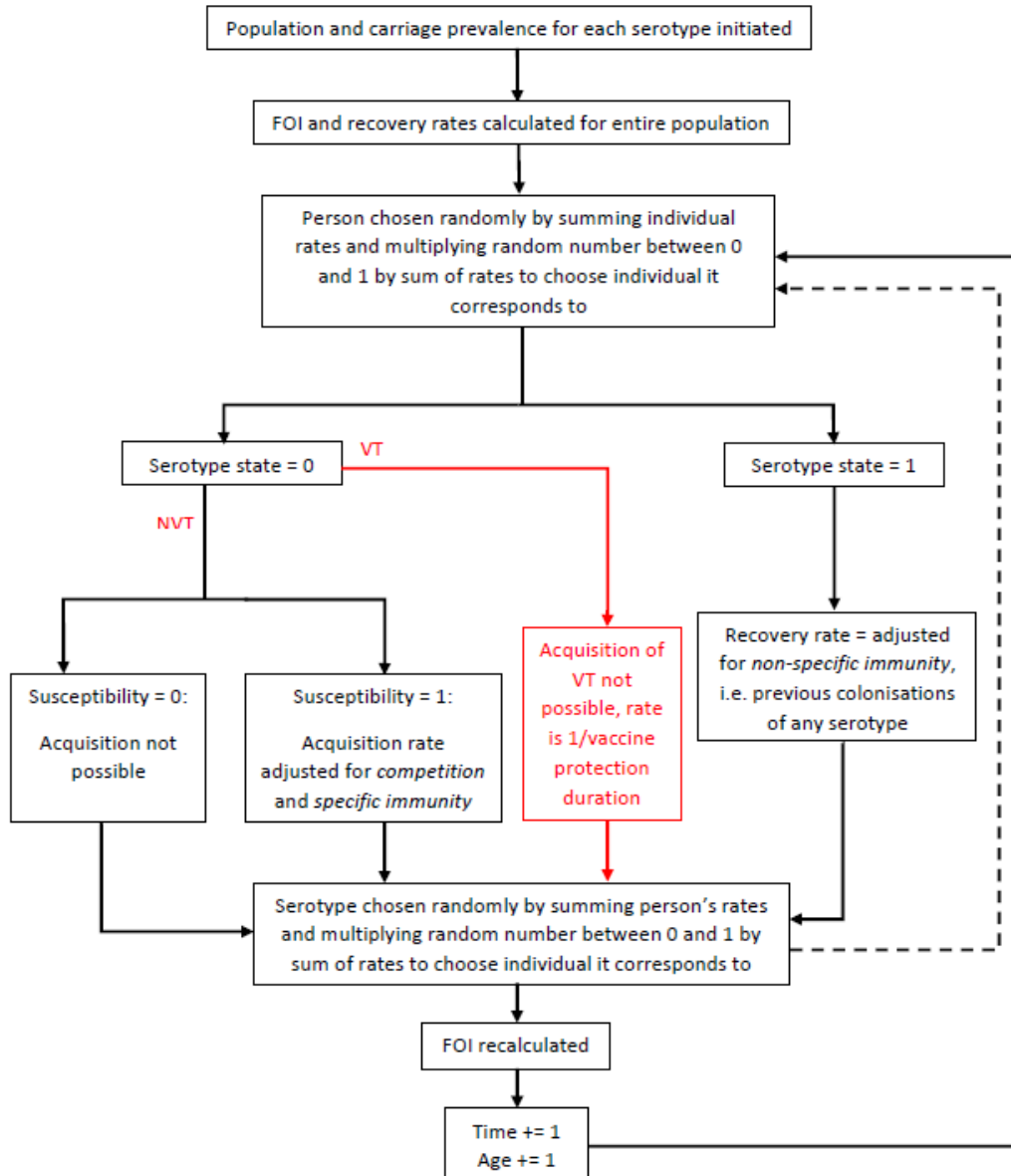
4. Design Concepts

Key features of the IBM are as follows:

Emergence: Transmission dynamics emerge from the process of immunity and serotype replacement. Serotype-specific immunity reduces the likelihood of re-acquisition of a certain serotype, enhancing serotype competition and potentially providing a mechanism to maintain serotype diversity. In this model, it is a reduction in the FOI for the following colonizing serotypes of that type. Non-specific immunity is a reduction in the duration of carriage. Additionally, as serotypes die out, they are replaced over time with new serotypes that emerge (serotype replacement). The number of disease cases are an emergent property of the model from the number of carriers and the serotype-specific invasiveness.

Adaptation: As individuals age, they are less likely to acquire carriage because of immunity. In response to vaccination, VT serotypes are eliminated from carriage, enabling NVT serotypes to become the more fit serotypes that cause disease.

362

Stochasticity: The choice of person to which the switching event happens is a random process drawn from a uniform distribution, as is the choice of the switching event.

Processes:
- Birth and mortality
- Serotype infection
- Serotype recovery

5. Observation
Parameters observed are the number of carriers and disease cases of each serotype following vaccination in the different scenarios.

6. Input
Longitudinal carriage prevalence data pre- and post-vaccination are used to fit the model. Previously obtained serotype-specific invasiveness estimates are used to convert carriers to disease cases.

7. Initialisation
A population of 5000 people is initialized with ages following the age demographics of the pertinent population and classified as children, adults or elderly. Each individual is characterized as either susceptible or carrier of each serotype, with a maximum carrying capacity of 2 serotypes, amounting to an initial carriage prevalence of 1% for each serotype.

8. Sub-models
Birth/mortality: individuals die when they reach the average life expectancy of the respective country and are replaced by an infant.
Infection/recovery: The model works as follows: first, for each serotype, the total FOI was calculated for the whole population. For each serotype, each individual was assigned either the FOI rate if they were susceptible or the recovery rate if they were carriers. To choose the person that the switching event happens to, a random number is drawn from a uniform distribution $x \in U (0,1)$ and multiplied by the sum of all individual rates. When compared to the cumulative sum of individual rates, this number will be in the interval of the rate corresponding to the chosen person. The chosen person will then have their serotype rates changed depending on how many serotypes they are carrying at the time they were chosen. If they are carrying 2 serotypes (i.e. the maximum carrying capacity), then all their other serotype FOI rates become 0 such that no other serotypes can be acquired. The maximum carrying capacity was chosen to be 2 to avoid complexity and because this is the most common number of serotypes detected in multi-serotype carriage [25]. The switching event happens to the serotype chosen using the same method as the person-choosing.
Immunity: Serotype-specific immunity reduces the FOI for the following colonizing serotypes of that type. Non-specific immunity is reduces the duration of carriage and is modelled as an exponential distribution.