Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

Olivier le Polain de Waroux

Thesis submitted in fulfilment of the requirement for the award of the degree of Doctor of Philosophy (PhD)

University of London

November 2016

Department of Infectious Disease Epidemiology,

Faculty of Epidemiology and Population Health,

London School of Hygiene and Tropical Medicine, London, United Kingdom

Declaration of own work

I, Olivier le Polain de Waroux, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Olivier le Polain de Waroux   Date: 18 November 2016

Abstract

Disease due to *Streptococcus pneumoniae* is a major cause of mortality and morbidity globally. Pneumococcal conjugate vaccines (PCVs) are being routinely introduced in immunisation programmes with support from Gavi, the Vaccine alliance, in low-income countries but uncertainty remains around the impact of different PCV introduction strategies in such settings, and in particular catch-up campaigns (CCs), which so far have not been conducted in Gavi-supported countries due to supply constraints.

This thesis explored the potential impact on nasopharyngeal carriage and disease of introducing PCV with and without CCs, in Nha Trang, Vietnam, through a dynamic transmission model. Vietnam is yet to introduce PCV vaccination, as are most South-East Asian countries. Additional studies on the vaccine efficacy against carriage (VE_C) and its waning, social contact patterns relevant for pneumococcal transmission, and age-specific epidemiology of carriage and serotype distribution were also conducted, the results of which fed directly into the transmission model.

A meta-analysis provided the first ever-global estimates of VE_C and its waning, by serotype and schedule. A large survey in southwest Uganda, collecting data on both social contact patterns and carriage from the same participants, shed light on which social contacts are important for pneumococcal transmission, showing that the frequency of close contacts, but not casual contacts, increased the colonisation risk, thereby informing the parameterisation of the transmission model. Results from the latter predicted elimination of vaccine type (VT) serotypes with near-complete replacement by non-VT across all age groups, within 10 years of PCV introduction with 90% coverage. The predicted benefit of CCs to expanded age groups was most pronounced in the first three years after PCV introduction, insofar that CCs result in limited introduction delays.

This thesis provided important insight into the epidemiology and transmission dynamics of pneumococci, to inform policy in countries that have not yet introduced PCV.

Acknowledgements

I would like to thank all colleagues and friends whom I have met, worked with, learned from and who have been great company over the last few years. Thanks to the whole modelling gang for making a jam-packed room 107 such an enviable place to work in. Many colleagues have contributed to this thesis in various ways, and I am very grateful to everyone with whom I have had the opportunity to work on this.

But in particular, I would mainly like to thank John Edmunds and Stefan Flasche, my thesis supervisors, for their support, scientific advice, their dynamism, motivation and for always being available – I mean it.

I would also like to thank the team at Epicentre, and in particular Sandra Cohuet, Fabienne Nackers, Juliet Amumpaire, Yap Boum II, Donny Ndazima, Celine Langendorf and Rebecca Grais for the opportunity to collaborate – it has been great. I am also very grateful to Francesco Checchi, who was instrumental in initiating the Uganda study, and would like to thank Anthony Scott for his input in this project too.

I am grateful to Kim Mulholland and Lay Myint Yoshida for their support with the Nha Trang project, as well as to Kensuka Takahashi, Koya Ariyoshi and Duc-Anh Dang.

Many thanks as well to David Goldblatt for his advice and guidance on several projects, David Prieto for his innovative suggestions to impromptu statistical questions, Adam Kucharski for his technical input in the contact study, and to Kate O'Brien and Maria Knoll for bringing their expertise and enthusiasm to the RESPICAR project. I am also grateful to Paul Heath and Jonathan Read for their review of the submitted thesis, insightful comments and helpful recommendations.

Finally, I am grateful to Sadia for her understanding – I am submitting the thesis, I promise.

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures

List of Supplementary Tables

List of Abbreviations

AMC	Advance Market Commitment
ARS	Acute respiratory symptoms
BIC	Bayesian Information Criterion
CC	Catch-up campaign
CC1	Catch-up campaign in <1 year olds
CC2	Catch-up campaign in <2 year olds
CC5	Catch-up campaign in <5 year olds
CI	Confidence Interval
CrI	Credible Interval
DIC	Deviance Information Criterion
DTP	Diphtheria-Tetanus-Pertussis
GMC	Geometric mean concentration
IPD	Invasive Pneumococcal Disease
LSHTM	London School of Hygiene and Tropical Medicine
MCMC	Markov Chain Monte Carlo
MUST	Mbarara University of Science and Technology
MSF	Médecins Sans Frontières
NT	Non-typeable
NVT	Non-vaccine type
OMPC	Outer membrane protein complex
OR	Odds ratio
PCV	Pneumococcal Conjugate Vaccine
PCV7	7-valent PCV
PCV10	10-valent PCV

PCV13	13-valent PCV
PP(S)V23	23-valent pneumococcal polysaccharide vaccine
R_0	Basic reproduction number
RSV	Respiratory Syncytial Virus
RV	Routine vaccination
SIS	Susceptible – Infected - Susceptible
VE_C	Vaccine efficacy against carriage acquisition
VE_{INV}	Vaccine efficacy against invasiveness
VE_{IPD}	Vaccine efficacy against IPD
VT	Vaccine type
WHO	World Health Organization

1. Introduction

1.1. Pneumococcal carriage and disease

Disease due to *Streptococus pneumoniae* (the pneumococcus) is a leading cause of morbidity and mortality, causing about 11% of deaths in children under five globally [1]. Community acquired pneumonia accounts for most of the severe pneumococcal disease burden worldwide, although the risk of severe morbidity and death is highest for invasive pneumococcal disease (IPD) such as bacteraemic pneumonia, meningitis and sepsis [1].

Carriage of *Streptococcus pneumoniae* in the nasopharynx is a precondition for pneumococcal disease, and the main reservoir for human-to-human transmission of the pathogen [2-4]. Pneumococci are part of the commensal nasopharyngeal flora, and as such, colonisation by pneumococci is both a frequent and generally asymptomatic event [5]. Infants generally become colonised within the first few weeks or months of life, and most children will have had at least one episode of nasopharyngeal colonisation in their first year of life [5-7]. Overall, the prevalence of carriage in children under five years of age is high, varying between settings from about 25% to 90% [8-10], and decreasing with age [10]. Colonisation rates tend to be higher in more deprived communities [11, 12], and are generally higher in developing than developed countries [8, 9]. Risk factors for carriage, other than age, include individual factors such as HIV infection, malnutrition or sickle cell disease [13-16], and community or ecological factors including seasonality [17, 18], pollution, inadequate hygiene, and increased social contact, particularly with and between young children. Several studies have shown that day-care centre attendance is associated with carriage prevalence among young children [19-23].

The extent to which some of the abovementioned factors explain differences in carriage prevalence rates between settings and countries remains insufficiently explored, but it is likely

that the sociodemographic structure and social mixing patterns are important factors explaining higher colonisation rates in low-income settings.

Although carriage is asymptomatic, it is a necessary precursor to disease; acquisition can occasionally lead to local mucosal disease, such as acute otitis media (AOM) or sinusitis, more rarely to lower respiratory tract infections and pneumonia, and in rare cases can lead to invasive disease [24]. The burden of pneumococcal disease tends to be disproportionally high in low income settings, not only because acquisition events tend to be more frequent, but also because underlying individual vulnerability factors such as malnutrition, HIV [13, 25, 26], sickle cell anaemia [27], as well as inadequate access to health services, can increase the risk of severe disease and death, particularly in infants and young children [28-32].

Pneumonia remains one of the main causes of mortality and morbidity in children under five years of age [1, 33]. In 2010, an estimated 1.4 million deaths due to pneumonia occurred globally, 43% of which were in Africa, and another 32% in South East Asia [33]. Although the aetiological ascertainment of radiological or clinical pneumonia remains difficult, *Streptococcus pneumoniae* is known to be one the main causative agents for both primary infection as well as secondary infection during upper respiratory viral illness such as measles or influenza [1, 26, 34]. Evidence from pneumococcal conjugate vaccine trials have consistently shown a reduction in the burden of radiological pneumonia among vaccinated children, with intervention studies from the Gambia [35, 36], South Africa [37], the Philippines [38], and among Navajo Indian Americans in the US [39] all showing substantial reductions in all-cause pneumonia among vaccinated arms, albeit of varying magnitude.

1.2.   Immune response to carriage and disease

Immunity resulting from natural exposure to *Streptococcus pneumoniae* in the nasopharynx is complex, and remains incompletely understood [40].

As nasopharyngeal carriage is a precondition to disease, and the main mechanism by which pneumococci are transmitted from person-to-person, elucidating the immune mechanisms against carriage is paramount. It is known that the local mucosal immune response plays an important role in clearance of acquired *S.pneumoniae*, and other pathogens of the commensal flora such as *moraxella catarrhalis*, *H influenzae* and others [24].

Carriage occurs through adherence of pneumococci to the epithelial lining of the nasopharynx, which is mediated by surface proteins, such as the pneumococcal surface protein A (PspA) or pneumococcal surface protein C (PspC) [24, 41]. It then requires the generation of inflammatory markers – as for example triggered during upper respiratory viral infections – to activate an inflammatory cascade that may result in the transcellular migration of pneumococci into the endothelium, resulting in invasive disease [24]. However, in the vast majority of cases, carriage acquisition remains asymptomatic, and the pathogen tends to be cleared within days or weeks.

There are nearly a 100 different known serotypes, which differ by the nature of their polysaccharide capsule and are thus immunologically distinct, even though common polysaccharides between certain serotypes results some degree of natural or vaccine-induces cross-protection (e.g. 6A and 6B). The polysaccharide capsule is the main virulence factor of pneumococci, as it allows the bacteria to evade phagocytosis, to some extent [5].

Differences in the nature of the polysaccharide likely explain some of the epidemiological characteristics of pneumococci in carriage and disease. Generally, serotypes with thicker capsules tend to escape opsonophagocytosis better than serotypes with thinner ones, and as such are also found more frequently in carriage, are carried for longer periods of time [42], and tend to predominate in carriage (and disease) among younger age groups. Epidemiological studies have consistently found that carriage decreases with age, and that the serotype distribution in carriage also tends to shift from being predominantly 'paediatric serotypes' in early childhood – which tend to be the most commonly found serotypes and are the serotypes targeted by the pneumococcal conjugate vaccines – to proportionally fewer of such serotypes as age increases [10]. Likewise, serotypes found in invasive disease tend to differ between age groups. For example, before the introduction of PCV, studies in England, Denmark and Germany found that serotype 1 and 5 accounted for about 20 – 25% of all serotypes isolated in IPD in neonates, whereas fewer than 5% of cases were due to such serotypes in older infants in the same studies [43-45]. Similarly, disease due to 'paediatric serotypes' tends to be proportionally much rarer amongst adults [46].

Those epidemiological findings support the evidence from laboratory and clinical studies that acquired anti-capsular (i.e. serotype-specific) immunity against nasopharyngeal colonisation, which requires the activation of the classic complement pathway, is one of the most important immune mechanisms against *S.pneumonaie* [47-49]. However, the exact mechanism underlying the protection against acquisition of carriage is still unclear and may involve memory B cells residing in the nasopharyngeal compartment responding to carriage and secreting local IgG or IgA rather than pre-existing circulating serum IgG. For example, studies have shown that both serotype-specific IgG and IgA antibodies can be found in the saliva of children colonised with a homologous serotype, further supporting a major role of a local serotype-specific immune reaction against pneumococcal polysaccharides as immune protection mechanism [50, 51].

Anti-capsular immunity is also central to the immune response to invasive infection. In particular, antibody binding to neutrophils, macrophages and other antigen presenting cells from the lung, liver or spleen, is crucial for opsonophagocytosis during invasive infections, and explains why asplenia, liver cirrhosis, or chronic splenic dysfunctions such as sickle cell anaemia tend to be risk factors for severe invasive infections[49].

In addition to antibody-mediated immunity against the capsule polysaccharide, there is now good evidence of acquired immunity to pneumococcal proteins, which are not common between serotypes [49, 52, 53]. Virolainen et al. [41] showed that children with lower antibody titres to PspA tended to be colonised more frequently with pneumococci than those with higher titres. Through a large household study in the UK, Goldblatt et al. [48] found that carriage induced a strong anticapsular response, but also a moderate antiprotein antibody response. As for anticapsular immunity, there is evidence that antibody-mediated immunity against pneumococcal surface proteins increase as children age [40].

In addition, there are also more recent studies suggesting that CD4+ T-helper 1 (Th1) and Th17 cells have an important role in clearance of carriage [54, 55]. Cell-mediated immunity is also thought to play an important factor in protecting against invasiveness, as suggested by studies showing that HIV positive individuals tend to be more prone to invasive infections, through deregulation of the T-cell response, before the CD4+ depletion [56].

In summary, both serotype-specific and non-specific immunity have important functions in the prevention against colonisation, and disease. Immunity from natural exposure to pneumococci can thus be summarised into three broad categories, namely (i) serotype-specific antibody-mediated immunity to the polysaccharide capsule of *Streptococcus pneumonia*, (ii) antibody-mediated immunity against pneumococcal surface proteins and (iii) cell-mediated antibody

independent immunity to pneumococcal proteins and the cell-wall polysaccharide, as described by Li et al. [40]. Antibody-mediated immunity will activate the classic complement route to opsonophagocytosis, whereas the cell-mediated immune mechanism includes CD4 cells that will activitate neutrophils and/or macrophages through the secretion of interleukin, and thereby kill *S.pneumonia* [40].

However, natural immunity is known to be imperfect, allowing for recurrent carriage episodes throughout life [42]. In young children in particular, immature immune systems tend to result in poor natural anti-capsular mediated immunity. For a few serotypes, such as serotype 14, anti-capsular immunity is known to provide good protection against re-acquisition (~90% for serotype 14), while for most other serotypes the antibody-mediated anti-capsular immunity is at best weak, hovering around 50% or less protection against reacquisition [57-59].

1.3. Pneumococcal conjugate vaccines

A pneumococcal polysaccharide vaccine covering 23 of at least 94 known *S.pneumoniae* serotypes (PPSV23) was licensed in 1983, and is still currently used particularly in the elderly and adult risk groups to prevent against invasive disease [60]. However, the protection conferred by polysaccharide vaccines in children is limited, due to poor immune responses to the polysaccharide antigens [61]. Conjugation of the capsular polysaccharide to a carrier protein enables the activation of T-cell mediated responses in children, and immunological memory. In addition, unlike PPSV23, pneumococcal conjugate vaccines provide mucosal immunity, also acting upon nasopharyngeal carriage acquisition, and thereby providing an indirect effect in addition to a direct community effect [62].

In contrast to natural immunity, vaccine-induced immunity through conjugate vaccines is reasonably good, with efficacies against carriage acquisition ranging by serotypes from about 40% to 80% [63], while the efficacy against invasive disease is higher, at about 90% overall [64].

In PCVs, the conjugation of capsule polysaccharides with protein carriers activates a T-cell help mechanism, which results in the generation of serotype-specific antibodies and B cell memory that much high protection against carriage acquisition than natural immunity [65, 66].

Nevertheless, although more immunogenic than natural immunity, PCV vaccination nonetheless requires multiple doses, particularly when given at a young age, and a booster dose >12 months of age to increase both the level and duration of protection [63]. Conjugate vaccines are also complex and costly to manufacture, given the different antigenic properties of each serotype. In addition, despite their success, problems of serotype replacement [67, 68], have brought researchers to explore alternative options to existing conjugate vaccines.

New vaccines in development target other immune mechanisms. These include targeting conserved surface protein antigens common to most (or all) pneumococci that are clinically relevant, in a serotype-independent way to provide broad cross serotype protection. A new range of vaccines are under development to target various proteins including pneumolysin, pneumococcal surface protein A (PspA) and protein C (PspC), pneumococcal surface antigen A (PsaA), neuraminidase enzymes, among others, most often as multi-component protein-based vaccine formulations [69-72].

The antibody-independent CD4+ T cell immune mechanism is also a target for vaccines in development, with the acapsular whole cell vaccine having now advanced through several stages of clinical trials [73].

A seven-valent pneumococcal conjugate vaccine (PCV7) was first introduced into the US childhood immunisation programme in 2000, and subsequently into the vaccination programmes of many other high-income countries. The introduction of PCV7 in low-income countries was limited by affordability issues (PCV are among the most expensive vaccines to manufacture) as well as a lower coverage by the vaccine of serotypes most commonly found in invasive disease in those countries [74]. For example, while in North America before PCV introduction more than 80% of all cases of IPD in children younger than five years were due to one of the seven serotypes included in PCV7 (4, 6B, 9V, 14, 18C, 19F, 23F), those serotypes only accounted for about 50% of IPD isolates in Africa and Asia [74].

Higher valency vaccines covering 10 serotypes (PCV10, Synflorix, manufactured by GlaxoSmithKline), and 13 serotypes (PCV13, Prevnar13, manufactured by Pfizer) have now replaced PCV7 in the routine vaccination schedule of most high-income settings. The expanded valency offers a better coverage of the main disease causative strains in developing countries, and both PCV10 and PCV13 have therefore steadily been introduced over the last few years in the routine immunisation programmes of low and lower-middle income countries, with support from Gavi, the Vaccine Alliance, through the pneumococcal Advance Market Commitment (AMC) [75, 76]. The three additional serotypes included in PCV10 (1, 5, 7F) are not very common in carriage, but can cause severe disease (particularly serotypes 1 and 5) with epidemic potential. PCV13 includes all PCV10 serotypes, in addition to serotypes 3, 6A and 19A, which are found commonly in both carriage and disease [9, 74, 77].

Pneumococcal conjugate vaccines prevent disease through two synergistic mechanisms; first, by providing some degree of protection against nasopharyngeal carriage acquisition of serotypes covered by the vaccine [63], and second, by preventing against further progression to disease, as a result of carriage [64], through serotype-specific antibody mediated immunity.

Through their effect on carriage, PCVs reduce the infection pool of targeted serotypes, thereby limiting the transmission of such serotypes in both vaccinated and unvaccinated individuals across age groups [78-81]. Such 'herd effect' has been observed in most high-income settings following PCV7 introduction [82, 83], and similar dynamics have been observed in developing countries, such as Kenya following PCV10 introduction [84], the Gambia following PCV7 and then PCV13 introduction [85]. Further evidence should be available from low-income settings in the coming years, given that at least 15 Gavi countries are currently conducting observational studies of PCV impact and effectiveness on disease and carriage endpoints [86].

One of the challenges posed by conjugate vaccines is that their impact on the transmission of the selected number of serotypes included in the vaccine formulation results in a (often proportional) increase in serotypes not covered by the vaccine, which fill the ecological niche vacated by vaccine-type (VT) serotypes [87-90]. Such 'serotype replacement' was first suggested by Lipsitch [90] in a simple illustrative mathematical model of colonising bacteria, where the potential replacement effects after vaccination were shown through the interplay of dyads of strains competing 'for space' and maintaining an equilibrium before vaccination; and where a rupture of this equilibrium through vaccination results in a lack of competition, and consequently an increased transmission of strains not targeted by the vaccine (i.e. 'replacement'). Replacement effects in nasopharyngeal carriage has been seen in most – if not all – settings following PCV introduction, for all types of commercially available pneumococcal conjugate vaccines, and in different types of epidemiological settings [68].

While serotype replacement has therefore tempered the impact of PCVs, the introduction of such vaccines in routine immunisation programmes has nevertheless resulted in a substantial reduction in the burden of pneumococcal disease, in both high and low income settings, due to the lower pathogenicity of replacing serotypes [83, 88, 89, 91-94].

However, despite consistent evidence from intervention studies and observational studies that the vaccine reduces severe disease, the magnitude of the effect varies across settings. For example, in a study evaluating the impact of PCV on IPD across 19 different surveillance sites from 13 different countries (including the USA, Canada, Australia, New Zealand, Great Britain, Czech Republic, Switzerland, Greece, Norway, Denmark, the Netherlands, Israel and Uruguay), the relative reduction in IPD incidence among <5 year olds of PCV7 introduction varied significantly between settings [95], some settings showing more than 50% decrease in the first year after PCV introduction whereas in other settings no change was yet observed; and relative reduction at 3 years varied from 24% to 83% reduction [95]. The routine introduction of the vaccine in high-income settings has also shown good evidence of herd protection against invasive disease, albeit of mixed magnitude [96]. While evidence of the impact on IPD in low-income settings is limited by the availability of robust surveillance capacity, evidence from surveillance settings such as in The Gambia and Kenya has consistently shown important reductions in VT carriage, and herd effects, following PCV introduction, but as in other settings, the magnitude of impact varies [85, 97, 98].

Differences in impact estimates between settings likely result from a combination of factors, including variations in vaccination coverage, vaccination schedules and introduction strategies, social mixing patterns and demographic composition, and baseline serotype composition in carriage and disease before PCV introduction. The latter in particular is the main predictor and likely driver of the long term impact of PCV [68]. In the shorter term, before the population impact of PCV reaches its new 'equilibrium' and before replacement effects are complete, factors such as the transmission rates in the population, itself linked to the frequency and age structure of social contacts, in addition to service delivery components such as vaccination

coverage and PCV introduction strategies, are likely to be key determinants of the speed with which population effects will be established, and resulting disease prevented.

1.4. Vaccination strategies, and problem statement

Routine childhood pneumococcal vaccination worldwide is a priority for WHO, particularly in countries where under five mortality is greater than 50 per 1,000 live births, most of which are in sub-Saharan Africa [99]. WHO currently recommends introducing PCV with either a '2+1' schedule, where two primary doses are administered in early infancy and a booster dose is given between the ages of 12 and 15 months, or with a '3+0' schedule, where three primary doses are provided before six months of age [100]. Although evidence suggests that the immune response obtained after a booster dose is higher than that obtained after the third primary dose [101], the epidemiological and long term impact of one or another schedule remains unclear. Currently, WHO recommends that schedules should be adapted to the setting's local epidemiological characteristics; a 3-dose primary schedule tends to be favoured over a 2-dose with a booster in countries where the burden of pneumococcal disease in infancy is particularly high and hence optimal protection in the first year is prioritised [100], with a first dose that can be administered as early as 6 weeks of age. As of December 2015, 54 countries had introduced PCV routinely, with four other introductions planned for 2016, including in Kyrgyzstan, Myanmar, Mongolia and Haiti [102].

Catch-up campaigns among older children at the introduction of PCV into childhood immunisation programmes can accelerate the impact of PCV, however there is no official recommendation from WHO about how to best implement those [100]. The benefits of catch-up campaigns include a more rapid establishment of direct vaccine effects on carriage and

disease on targeted cohorts, but also potentially a faster way for indirect population (i.e. 'herd') effects of the vaccine to settle in, due to a larger age cohort of vaccinated children having social contacts with unvaccinated individuals in the community [103]. While it is obvious that targeting more age groups through campaigns will lead to a more rapid population effect of the vaccine, potential additional health gains need to be balanced against health service, resource, and financial constraints. As of October 2016, at least 54 countries had introduced PCV with Gavi's support, and with the exception of a few surveillance and research settings (such as in Kilifi, Kenya [104], or the Basse region in The Gambia [85]), no nationwide catch-up campaigns at vaccine introduction have been conducted in countries supported by Gavi, due to resource and supply constraints [102]. In addition, decisions about catch-up strategies may have been delayed by the limited evidence on the cost benefit of campaigns in such settings, as well as supply constraints, the potential disruptive effect of campaigns on routine services, and remaining questions on how to best optimise campaigns, and maximise their impact on disease with the resources and vaccine supplies available.

Although post-PCV surveillance data exist for countries that have introduced PCV with a catch-up campaign, such as the UK [105] or Uruguay [106], there is no population-based study providing direct comparison of the differential impact of various strategies. Such comparisons could be undertaken through large intervention studies in settings that have not yet introduced the vaccine, but require time, resources and sophisticated epidemiological design. In the absence of evidence from intervention studies, modelling approaches can be used to explore potential impact of different vaccination options, and guide related policy questions.

This thesis was driven by two initial key and interlinked questions related to the pneumococcal catch-up strategies at the introduction of PCV in a Gavi-eligible country, which were approached through a dynamic model of pneumococcal transmission. The questions were as follows:

1) First, compared to a routine introduction of the vaccine, what is the likely impact on carriage and disease of catch-up strategies targeting infants only (<1 year), young children (<2 years) or all children <5 years?

2) Second, given supply constraints, could a delayed introduction of PCV to allow for a catch-up campaign have any benefit over a timelier introduction of PCV in a routine-only vaccination strategy?

These questions are particularly timely in the current policy context. Last year, with the exception of Nigeria, vaccine supplies have increased to now meet demand in all other Gavi countries, easing off some of the supply constraints and enabling Gavi and eligible countries to consider PCV introduction plans that would include catch-up campaigns without compromising the introduction to other settings or existing programmes [102]. Research to establish the best use of vaccines in high burden settings is also high on WHO's agenda [107].

While modelling approaches may provide further insight into the potential short, medium and long term impact of vaccination options, models crucially depend on the robustness and validity of assumptions made about how pathogens are transmitted between humans, what contact events are important for transmission, and the structure of effective interpersonal encounters in the population [108-110], and, in the context of pneumococcal dynamics, what the characteristics are of carriage and serotype distribution across population age groups before the vaccine is introduced. Other central questions in the exploration of alternative pneumococcal strategies through modelling include how well the vaccine protects against carriage and disease, for primary, booster or catch-up doses, and how long the vaccine would protect for. Although carriage is a key determinant of the impact of the vaccine in a defined population, the majority of PCV trials have overlooked the impact of the vaccine on

colonisation, or have considered the impact on carriage on small study samples and over a short period of time, particularly given that the impact of PCV on nasopharyngeal carriage was not considered in the licensure process [111].

Hence, a number of questions around the consolidation of our knowledge about PCV efficacy on carriage acquisition, dynamics of carriage in the community, and how pneumococci are transmitted in the community underpinned the work of this thesis, also feeding directly into the structure and parameterisation of the transmission model.

1.5. Dynamic transmission models of *Streptococcus pneumoniae*

Dynamic mathematical models of infectious disease transmission are increasingly used to better understand underlying mechanisms and factors contributing to transmission, predict disease trends, and disease control options and strategies, including vaccination [108].

Transmission models are, by definition, a simplification of a complex interplay between ecological and human population processes. The structure of a model and its degree of sophistication will depend, in part, on the questions to address.

Lipsitch designed one of the first transmission model for colonising bacteria such as *Streptococcus pneumoniae* [90], drawing on the experience from previous models of between-strain interaction [112] and superinfection [113, 114]. Lipsitch's compartmental model, initially for two colonising (and mutually competing) serotypes and later extended to multiple serotypes, was designed to explore the potential impact of between-serotype competition on serotype replacement when introducing serotype-specific vaccines. In this model, individuals were assumed to be in one of four compartments of carriage state namely (i) susceptible to

30

colonisation by any strain, (ii) colonised by one strain, (iii) colonised by another strain, or (iv) co-colonised; and reverting back to a susceptible state after carriage assuming no natural immunity (i.e. a 'Susceptible – Infected – Susceptible' (SIS) model). The model was designed more as an illustrative piece of work, exploring the potential impact on carriage and disease dynamics of introducing serotype-specific vaccines under different assumptions about between-serotype competition, vaccine efficacy and vaccine types. While the model has inherent limitations and simplifications, it laid out the basis for dynamic models based on a similar 'diamond-shaped' structure, which are discussed hereafter.

In line with Lipsitch's model structure, Choi et al. [115] and Melegaro et al. [116] designed deterministic compartmental models to predict the long term impact on carriage and disease of introducing PCV7 [115, 116] as well as the differential impact on IPD of switching to PCV13 after several years of PCV7 use [117]. Melegaro et al. [116] first developed an age-specific deterministic SIS model of carriage acquisition and clearance, allowing for co-infection. In Melegaro's model serotypes are not modelled individually but are pooled instead, with the group of vaccine type (VT) serotypes modelled jointly and separately from the group of non-vaccine type (NVT) serotypes, although the model allows for co-colonisation with one VT and one NVT serotype. Vaccination is introduced in the model as a factor reducing the force of infection of VT serotypes among vaccinated individuals, thereby creating two levels of vaccine-induced protection (vaccinated or not) for each of the four model compartments (susceptible, VT colonisation, NVT colonisation and VT-NVT dual colonisation). The model was fitted to carriage and disease data and long term IPD trends were predicted based on estimations of vaccine efficacy, competition between VT and NVT serotypes, and assumptions about duration of carriage, vaccination coverage and the stability of the pathogenicity of the group of VT and NVT serotypes over time. The model was later expanded to better account for the dose-specific efficacy of PCV, with weaker and shorter-lived vaccine-immune protection conferred by two

primary infant doses and stronger and longer protection acquired after the booster dose [115]. Choi et al. [117] also further expanded the model to specifically explore the impact of replacing PCV7 by PCV13 in the UK's routine immunization programme, and therefore included the group of the six PCV13 serotypes not included in PCV7 as a third group of serotypes in the model structure. A similar model structure was used by De Cao et al. [118] in a model applied to the Netherlands, to predict the long-term impact of PCV13 on carriage and disease.

In all models where serotypes are grouped, replacement effects result from competition parameters between such groups (i.e. a reduction in the likelihood of acquisition in the presence of carriage with another serotype in the group), which in the models described above is dichotomised between carriers of serotypes that are either vaccine serotypes or serotypes not targeted by the vaccine. Although competition between groups of serotypes rather than single strain competition moved these models away from more realistic biological processes (i.e. different serotypes compete between themselves regardless of their inclusion or not in the vaccine), such deterministic models have performed well at capturing the overall population-wide pneumococcal dynamics post vaccine introduction and as such, are useful in helping plan and optimise vaccine introduction strategies. While individually modelling each of the nearly 100 serotypes may not be particularly useful, given the very low prevalence of some and lack of complete data in most cases, dividing groups of serotypes into smaller distinct groups of serotypes – for example by pathogenicity of prevalence – may more closely represent the transmission dynamics, although the optimal number of subgroups remains unknown.

In a model applied to the Gambia, Bottomley et al. [119] categorised serotypes into six different classes, based on the serotype's pre-PCV prevalence, with three classes for groups of NVT serotypes of different carriage prevalence (high, medium and low) and three groups of VT serotypes (of high, medium and low prevalence In that model, serotype coexistence results from a proportion of colonised individuals gaining lifelong immunity. Results from Bottomley

et al. predicted a rapid decline in the carriage prevalence of vaccine serotypes of low and medium pre-PCV prevalence, incomplete reductions at 10 years of the prevalence of VT serotypes of high prevalence before PCV introduction, and little overall change in carriage prevalence due to serotype replacement. Nurhonen et al. [120] developed an individual contact network model, with serotypes divided into 11 classes, based on a combination of characteristics of invasiveness, carriage prevalence and inclusion or not in the vaccine formulation, and applied to a population with the demographic and epidemiological characteristics of Finland. Unlike the 'diamond shaped' models previously mentioned[115, 118], in Nurhonen's microsimulation model replacement by NVT is not proportionally to their prevalence before PCV introduction, but rather to the fitness of non-vaccine serotypes. However, the scientific gain from such modelling approach remains to be evaluated. Results from Nurhonen et al. [120] were broadly in line with the findings from simpler deterministic models described earlier, and in concordance with epidemiological studies post PCV introduction, showing near-elimination of vaccine-type serotypes within 5 – 10 years after vaccine introduction, with a vaccination coverage of 90%, and assuming a vaccine efficacy against carriage acquisition of 50%.

Other dynamic models have also been used to predict and understand post-vaccination trends, conduct economic evaluations and inform vaccination policy. Van Effelterre et al. [121] used a deterministic model to explore and compare the relative contribution of PCV7 and antibiotic pressure to the increase in serotype 19A IPD cases observed among <2y olds in the USA among following PCV7 implementation. The model resembles those described earlier [115, 116], although it specifically includes antibiotic treatment with compartments of colonised and non-colonised children with and without treatment. Wu et al. [122] applied a dynamic SIS transmission model to data from Taiwan to undertake an economic evaluation of the impact of introducing PCV13. Zhang et al. [123] used a dynamic deterministic model to help understand

the influence of temporary immunity and prevalence levels on competition. Here, unlike previous SIS models a compartment of temporary resistance to carriage acquisition (R) was added (SIRS model) to better understand the influence of temporary immunity to colonisation on prevalence levels and ecological competition between serotypes.

Models applied to communities in the USA and in Sweden have also been used to explore the relative contribution of different social factors on the transmission of *S.pneumoniae*, including household size, day care attendance and class size [124, 125].

In the context of PCV vaccination policy, models such as those described earlier can be used to add to the evidence from trials and observational studies, and allow one to address questions that may be difficult to address through other epidemiological approaches for reasons such as a lack of follow-up time, difficulties of comparing the population impact of many different introduction strategies in observational trials due epidemiological, logistical and programmatic challenges, or where intervention studies may be unethical, such as exploring the impact of PCV catch-up campaigns in a context of low vaccine uptake. In addition, models provide a tool to explore the longer-term impact of interventions, and help policy-makers appraise the different policy options based on medium- or long-term impact estimates, also feeding into economic evaluations.

However, the inherent (and purposely) simplification of complex transmission dynamics and the departure – to some extent – from pathogen dynamics by pooling all serotypes into defined groups, make the abovementioned models less suitable to understand underlying biological and immunological mechanisms explaining the observed epidemiological patterns.

Markov transition models were applied to longitudinal carriage data from Finland [126], Denmark [2], Kenya [127] and the UK [128], to estimate clearance and competition parameters

between serotypes in a steady-state situation, and transition models were used to model acquisition rates for multiple colonisation among families in Bangladesh [126].

More complex individual-based models such as those developed by Cobey and Lipsitch [42] as well as by Flasche et al. [129] have been developed to explore underlying mechanisms to the ecological and epidemiological patterns of *S.pneumoniae* carriage. Such theoretical and mostly data-free modelling approaches shed light on the joint contribution of serotype-specific and nonspecific immunity on serotype competition and coexistence, and their implications for the impact of pneumococcal vaccines. Both models here assume some level of serotype-specific and nonspecific immunity following colonisation. Although assumptions made about level and duration of immunity differed between studies, both studies reached similar conclusions about the key contribution of immunological factors to epidemiological patterns, including serotype coexistence, prevalence of carriage and the emergence of epidemic strains, and their implication for vaccine development and policy. Interestingly, a theoretic finding from Flasche et al.'s work [129] is the synergistic effect of competition and selective vaccination in eliminating targeted serotypes when such vaccines would otherwise fail to reach the herd immunity threshold. Cobey and Lipsitch's model [42] showed that reductions in nonspecific immunity as a result of carriage reduction following vaccination could lead to the re-emergence of previously eliminated serotypes. Compared to the deterministic models mentioned earlier [115-117, 121-123, 130], the advantage of individual-based serotype-specific models is to better explore the biological mechanisms of acquisition and transmission, and to appraise the contribution of several underlying factors to the epidemiological and ecological patterns of *Streptococcus pneumoniae*.

In the context of the roll out of PCV in many developing countries, carefully designed mathematical models could offer new insights into the impact of the vaccine on carriage and

disease across all age groups and in particular explore the differential impact on transmission, morbidity and mortality of various vaccination scenarios, including catch-up campaigns. As such, models therefore add onto the evidence from existing or planned observational or interventional studies, or offer an alternative to such studies when design, financial, logistical and ethical challenges may limit their implementation.

1.6.   Aims and objectives

1.6.1.   Main aim

The main aim of this thesis is to study pre-vaccination transmission dynamics of *Streptococcus pneumoniae* in a Gavi-support eligible setting, predict the likely impact on carriage and disease of routine vaccination, and explore the differential impact of catch-up campaigns at PCV introduction on carriage and invasive pneumococcal disease dynamics across age groups.

1.6.2.   Specific objectives, and structure of the thesis

For this to be achieved, the thesis is articulated around the following main objectives:

1. To undertake a systematic review and meta-analysis of the prevalence and serotype distribution of *S.pneumoniae* in carriage, for groups of vaccine types and non-vaccine serotypes, such as to predict the prevalence in older age groups as a function of that in children under five years of age, and parameterise the model for Nha Trang (Vietnam) accordingly.

2. To estimate the vaccine efficacy against carriage acquisition, and the duration of protection against carriage, through a meta-regression model of available trial data, in order to parameterise the model applied to the population of Nha Trang.

3. To undertake a study of social mixing patterns and of *S.pneumoniae* carriage in South West Uganda, in order to determine whether particular contacts and transmission patterns are associated with *S.pneumoniae* acquisition.

4. To use the evidence generated in the first three specific objectives to develop and parameterise a dynamic transmission model of *S.pneumoniae*, and explore two specific questions, namely, (i) what is the differential impact on carriage and disease of introducing PCV with and without catch-up campaigns targeting three different age groups (<1 years, <2 years, <5years) and (ii) under resource constraints, how would introduction delays with a catch-up campaign impact on carriage and disease compared to a more timely introduction of PCV routinely, without such campaign.

1.7. Structure of the thesis

The abovementioned objectives were addressed through specific independent research projects, with different collaborators and study groups, and were conducted in the order described above.

The results have been published (or submitted) in five different manuscripts, which form the five main results chapters of this thesis. The third objective was divided into two separate chapters; one about the details of social mixing patterns in this setting, and the other one exploring the type of social contacts associated with pneumococcal carriage.

In addition to the published papers, some sections have been expanded with appendices to expand on aspects from the methods or results which could not be addressed in details in each publication.

The differential impact of catch-up campaigns was studied through a model applied to the population of Nha Trang in Vietnam. Few Asian countries have introduced PCV; in 2012 Pakistan was the first Asian country to routinely vaccinate infants with PCV, followed by Nepal, Cambodia and Lao PDR [102]. Vietnam is one of seven countries that has not yet applied for Gavi's support through AMC due to a mix of sustainability concerns and prioritisation of rotavirus vaccine but is eligible to do so [102]. Work in Vietnam has therefore direct relevance for local vaccine policy, and could inform introduction strategies both in Vietnam, and in other neighbouring countries with similar epidemiology and transmission dynamics. Absence of colonisation data in adults and elderly, and the lack of post-vaccination data for this setting motivated the first two chapters of the thesis, namely (i) to better understand (and predict) the prevalence of carriage and serotype distribution in older age groups based on carriage estimates in children under five years of age and (ii) obtain robust estimates of the efficacy and duration of protection of PCV against nasopharyngeal carriage acquisition.

Objective 3 was addressed through a field survey in Uganda, which included a nasopharyngeal carriage survey, and a study of social contacts among a subsample of the same study participants. The combination of carriage data and social contact data made it possible to explore in finer details how individual social behaviour shapes individual-level risk of *S.pneumoniae*, and thereby what social contacts are important to include in pneumococcal transmission models.

Additional work on pneumococcal dynamics and vaccine optimisation will follow from the study in Uganda, but is beyond the scope of this thesis (see Chapter 7).

38

References

1.   O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet. 2009;374(9693):893-902. doi: 10.1016/S0140-6736(09)61204-6. PubMed PMID: 19748398.

2.   Auranen K, Mehtala J, Tanskanen A, Kaltoft S. Between-strain competition in acquisition and clearance of pneumococcal carriage--epidemiologic evidence from a longitudinal study of day-care children. AmJEpidemiol. 2010;171(2):169-76. doi: kwp351 [pii];10.1093/aje/kwp351 [doi].

3.   Auranen K, Rinta-Kokko H, Goldblatt D, Nohynek H, O'Brien KL, Satzke C, et al. Design questions for Streptococcus pneumoniae vaccine trials with a colonisation endpoint. (Special Section: Nasopharyngeal carriage in the pneumococcal vaccine licensure process.). Vaccine. 2014;32(1):159-64.

4.   Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL, et al. The fundamental link between pneumococcal carriage and disease. Expert review of vaccines. 2012;11(7):841-55. doi: 10.1586/erv.12.53. PubMed PMID: 22913260.

5.   Bogaert D, Belkum Av, Sluijter M, Luijendijk A, Groot Rd, Rumke HC, et al. Colonisation by Streptococcus pneumoniae and Staphylococcus aureus in healthy children. Lancet. 2004;363(9424):1871-2.

6.   Abdullahi O, Karani A, Tigoi CC, Mugo D, Kungu S, Wanjiru E, et al. The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. PloS one. 2012;7(2):e30787. Epub 2012/03/01. doi: 10.1371/journal.pone.0030787. PubMed PMID: 22363489; PubMed Central PMCID: PMC3282706.

7.   Turner P, Turner C, Jankhot A, Helen N, Lee SJ, Day NP, et al. A longitudinal study of Streptococcus pneumoniae carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. PloS one. 2012;7(5).

8.   Usuf E, Bottomley C, Adegbola RA, Hall A. Pneumococcal carriage in sub-Saharan Africa--a systematic review. PloS one. 2014;9(1):e85001. doi: 10.1371/journal.pone.0085001. PubMed PMID: 24465464; PubMed Central PMCID: PMC3896352.

9.   Adegbola RA, DeAntonio R, Hill PC, Roca A, Usuf E, Hoet B, et al. Carriage of Streptococcus pneumoniae and Other Respiratory Bacterial Pathogens in Low and Lower-Middle Income Countries: A Systematic Review and Meta-Analysis. PloS one. 2014;9(8):e103293. doi: 10.1371/journal.pone.0103293. PubMed PMID: 25084351; PubMed Central PMCID: PMC4118866.

10.   Le Polain de Waroux O, Flasche S, Prieto-Merino D, Edmunds WJ. Age-Dependent Prevalence of Nasopharyngeal Carriage of Streptococcus pneumoniae before Conjugate Vaccine Introduction: A Prediction Model Based on a Meta-Analysis. PloS one. 2014;9(1):e86136. doi: 10.1371/journal.pone.0086136. PubMed PMID: 24465920; PubMed Central PMCID: PMC3900487.

11.   Huang SS, Finkelstein JA, Rifas-Shiman SL, Kleinman K, Platt R. Community-level predictors of pneumococcal carriage and resistance in young children. American journal of epidemiology. 2004;159(7):645-54. PubMed PMID: 15033642.

12.   Mackenzie GA, Leach AJ, Carapetis JR, Fisher J, Morris PS. 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 infectious diseases. 2010;10(304).

13.   Heinsbroek E, Tafatatha T, Phiri A, Ngwira B, Crampin AC, Read JM, et al. Persisting high prevalence of pneumococcal carriage among HIV-infected adults receiving antiretroviral therapy in Malawi: a cohort study. AIDS. 2015;29(14):1837-44. doi:

10.1097/QAD.000000000000755. PubMed PMID: 26218599; PubMed Central PMCID: PMCPMC4568891.

14.   Conklin LM, Bigogo G, Jagero G, Hampton L, Junghae M, da Gloria Carvalho M, et al. High Streptococcus pneumoniae colonization prevalence among HIV-infected Kenyan parents in the year before pneumococcal conjugate vaccine introduction. BMC infectious diseases. 2016;16:18. doi: 10.1186/s12879-015-1312-2. PubMed PMID: 26774803; PubMed Central PMCID: PMCPMC4715316.

15.   Verhagen LM, Gomez-Castellano K, Snelders E, Rivera-Olivero I, Pocaterra L, Melchers WJ, et al. Respiratory infections in Enepa Amerindians are related to malnutrition and Streptococcus pneumoniae carriage. The Journal of infection. 2013;67(4):273-81. doi: 10.1016/j.jinf.2013.06.010. PubMed PMID: 23796866.

16.   Rivera-Olivero IA, Nogal Bd, Fuentes M, Cortez R, Bogaert D, Hermans PWM, et al. Immunogenicity of a 7-valent pneumococcal conjugate vaccine (PCV7) and impact on carriage in Venezuelan children at risk of invasive pneumococcal diseases. Vaccine. 2014;32(31):4006-11.

17.   Numminen E, Chewapreecha C, Turner C, Goldblatt D, Nosten F, Bentley SD, et al. Climate induces seasonality in pneumococcal transmission. Sci Rep. 2015;5:11344. doi: 10.1038/srep11344. PubMed PMID: 26067932; PubMed Central PMCID: PMCPMC4464306.

18.   Bojang A, Jafali J, Egere UE, Hill PC, Antonio M, Jeffries D, et al. Seasonality of Pneumococcal Nasopharyngeal Carriage in Rural Gambia Determined within the Context of a Cluster Randomized Pneumococcal Vaccine Trial. PloS one. 2015;10(7):e0129649. doi: 10.1371/journal.pone.0129649. PubMed PMID: 26132206; PubMed Central PMCID: PMCPMC4488590.

19.   Ahl J, Melander E, Odenholt I, Tvetman L, Thornblad T, Riesbeck K, et al. Risk factors for pneumococcal carriage in day care centers: a retrospective study during a 10-year period. Pediatric Infectious Disease Journal. 2014;33(5):536-8.

20.   Hsieh Y, Chiu C, Chang K, Huang Y, Chen C, Kuo C, et al. The impact of the heptavalent pneumococcal conjugate vaccine on risk factors for streptococcus pneumoniae carriage in children. Pediatric Infectious Disease Journal. 2012;31(9):e163-e8.

21.   Givon-Lavi N, Fraser D, Porat N, Dagan R. Spread of Streptococcus pneumoniae and antibiotic-resistant S. pneumoniae from day-care center attendees to their younger siblings. The Journal of infectious diseases. 2002;186(11):1608-14. doi: 10.1086/345556. PubMed PMID: 12447737.

22.   Dagan R, O'Brien KL. Modeling the association between pneumococcal carriage and child-care center attendance. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2005;40(9):1223-6. doi: 10.1086/428585. PubMed PMID: 15825021.

23.   Toledo ME, Casanova MF, Linares-Perez N, Garcia-Rivera D, Peraza GT, Pina IB, et al. Prevalence of Pneumococcal Nasopharyngeal Carriage Among Children 2-18 Months of Age: Baseline Study Pre-Introduction of Pneumococcal Vaccination in Cuba. The Pediatric infectious disease journal. 2016. doi: 10.1097/INF.0000000000001341. PubMed PMID: 27649366.

24.   Bogaert D, De Groot R, Hermans PW. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. The Lancet infectious diseases. 2004;4(3):144-54. doi: 10.1016/S1473-3099(04)00938-7. PubMed PMID: 14998500.

25.   Heinsbroek E, Tafatatha T, Chisambo C, Phiri A, Mwiba O, Ngwira B, et al. Pneumococcal Acquisition Among Infants Exposed to HIV in Rural Malawi: A Longitudinal Household Study. American journal of epidemiology. 2016;183(1):70-8. doi: 10.1093/aje/kwv134. PubMed PMID: 26628514; PubMed Central PMCID: PMCPMC4690474.

26.   Albrich WC, Pride MW, Madhi SA, Callahan J, Adrian PV, French R, et al. Multiplex urinary antigen detection for 13 Streptococcus pneumoniae serotypes improves diagnosis of pneumococcal pneumonia in South African HIV-infected adults. Journal of clinical microbiology. 2016. doi: 10.1128/JCM.01573-16. PubMed PMID: 27847374.

27.   Rivera-Olivero IA, Del Nogal B, Fuentes M, Cortez R, Bogaert D, Hermans PW, et al. Immunogenicity of a 7-valent pneumococcal conjugate vaccine (PCV7) and impact on carriage in Venezuelan children at risk of invasive pneumococcal diseases. Vaccine. 2014;32(31):4006-11. doi: 10.1016/j.vaccine.2014.04.027. PubMed PMID: 24837505.

28.   Cohen AL, Hyde TB, Verani J, Watkins M. Integrating pneumonia prevention and treatment interventions with immunization services in resource-poor countries. Bulletin of the World Health Organization. 2012;90(4):289-94. doi: 10.2471/BLT.11.094029. PubMed PMID: 22511825; PubMed Central PMCID: PMC3324866.

29.   Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375(9730):1969-87. doi: 10.1016/S0140-6736(10)60549-1. PubMed PMID: 20466419.

30.   Global Action Plan for prevention and control of pneumonia (GAPP) Geneva: World Health Organization; 2008. Available from: http://whqlibdoc.who.int/hq/2009/WHO_FCH_CAH_NCH_09.04_eng.pdf.

31.   Jordano Q, Falco V, Almirante B, Planes AM, del Valle O, Ribera E, et al. Invasive pneumococcal disease in patients infected with HIV: still a threat in the era of highly active antiretroviral therapy. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2004;38(11):1623-8. doi: 10.1086/420933. PubMed PMID: 15156452.

32.   Feikin DR, Feldman C, Schuchat A, Janoff EN. Global strategies to prevent bacterial pneumonia in adults with HIV disease. The Lancet infectious diseases. 2004;4(7):445-55. doi: 10.1016/S1473-3099(04)01060-6. PubMed PMID: 15219555.

33.   Liu L, Johnson HL, Cousens S, Perin J, Scott S, Lawn JE, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. Lancet. 2012;379(9832):2151-61. doi: 10.1016/S0140-6736(12)60560-1. PubMed PMID: 22579125.

34.   Madhi SA, De Wals P, Grijalva CG, Grimwood K, Grossman R, Ishiwada N, et al. The burden of childhood pneumonia in the developed world: a review of the literature. The Pediatric infectious disease journal. 2013;32(3):e119-27. doi: 10.1097/INF.0b013e3182784b26. PubMed PMID: 23099423.

35.   Cutts FT, Zaman SM, Enwere G, Jaffar S, Levine OS, Okoko JB, et al. Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. Lancet. 2005;365(9465):1139-46. doi: 10.1016/S0140-6736(05)71876-6. PubMed PMID: 15794968.

36.   Mackenzie GA, Bottomley C, van Hoek AJ, Jeffries D, Ota M, Zaman SM, et al. Efficacy of different pneumococcal conjugate vaccine schedules against pneumonia, hospitalisation, and mortality: re-analysis of a randomised trial in the Gambia. Vaccine. 2014;32(21):2493-500. doi: 10.1016/j.vaccine.2014.02.081. PubMed PMID: 24631086.

37.   Klugman KP, Madhi SA, Huebner RE, Kohberger R, Mbelle N, Pierce N. A trial of a 9-valent pneumococcal conjugate vaccine in children with and those without HIV infection. The New England journal of medicine. 2003;349(14):1341-8. Epub 2003/10/03. doi: 10.1056/NEJMoa035060

349/14/1341 [pii]. PubMed PMID: 14523142.

38.   Lucero MG, Nohynek H, Williams G, Tallo V, Simoes EA, Lupisan S, et al. Efficacy of an 11-valent pneumococcal conjugate vaccine against radiologically confirmed pneumonia among children less than 2 years of age in the Philippines: a randomized, double-blind, placebo-controlled trial. The Pediatric infectious disease journal. 2009;28(6):455-62. doi: 10.1097/INF.0b013e31819637af. PubMed PMID: 19483514.

39.   O'Brien KL, Moulton LH, Reid R, Weatherholtz R, Oski J, Brown L, et al. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group

randomised trial. Lancet. 2003;362(9381):355-61. doi: 10.1016/S0140-6736(03)14022-6. PubMed PMID: 12907008.

40.   Li Y, Gierahn T, Thompson CM, Trzcinski K, Ford CB, Croucher N, et al. Distinct effects on diversifying selection by two mechanisms of immunity against Streptococcus pneumoniae. PLoS pathogens. 2012;8(11):e1002989. doi: 10.1371/journal.ppat.1002989. PubMed PMID: 23144610; PubMed Central PMCID: PMCPMC3493470.

41.   Virolainen A, Russell W, Crain MJ, Rapola S, Kayhty H, Briles DE. Human antibodies to pneumococcal surface protein A in health and disease. The Pediatric infectious disease journal. 2000;19(2):134-8. PubMed PMID: 10694000.

42.   Cobey S, Lipsitch M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. Science. 2012;335(6074):1376-80. Epub 2012/03/03. doi: 10.1126/science.1215947. PubMed PMID: 22383809; PubMed Central PMCID: PMC3341938.

43.   von Kries R, Hermann M, Hachmeister A, Siedler A, Schmitt HJ, Al-Lahham A, et al. Prediction of the potential benefit of different pneumococcal conjugate vaccines on invasive pneumococcal disease in German children. The Pediatric infectious disease journal. 2002;21(11):1017-23. doi: 10.1097/01.inf.0000036013.24857.a2. PubMed PMID: 12442022.

44.   Kaltoft, Zeuthen N, Konradsen HB. Epidemiology of invasive pneumococcal infections in children aged 0-6 years in Denmark: a 19-year nationwide surveillance study. Acta Paediatr Suppl. 2000;89(435):3-10. PubMed PMID: 11194794.

45.   Sleeman K, Knox K, George R, Miller E, Waight P, Griffiths D, et al. Invasive pneumococcal disease in England and Wales: vaccination implications. The Journal of infectious diseases. 2001;183(2):239-46. doi: 10.1086/317924. PubMed PMID: 11120930.

46.   Hausdorff WP, Feikin DR, Klugman KP. Epidemiological differences among pneumococcal serotypes. Lancet Infectious Diseases. 2005;5(2):83-93.

47.   Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, Lipsitch M. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. The Journal of infectious diseases. 2008;197(11):1511-8. doi: 10.1086/587941. PubMed PMID: 18471062.

48.   Goldblatt D, Hussain M, Andrews N, Ashton L, Virta C, Melegaro A, et al. Antibody responses to nasopharyngeal carriage of Streptococcus pneumoniae in adults: a longitudinal household study. JInfectDis. 2005;192(3):387-93. doi: JID34243 [pii];10.1086/431524 [doi].

49.   Malley R. Antibody and cell-mediated immunity to Streptococcus pneumoniae: implications for vaccine development. J Mol Med (Berl). 2010;88(2):135-42. doi: 10.1007/s00109-009-0579-4. PubMed PMID: 20049411.

50.   Simell B, Korkeila M, Pursiainen H, Kilpi TM, Kayhty H. Pneumococcal carriage and otitis media induce salivary antibodies to pneumococcal surface adhesin A, pneumolysin, and pneumococcal surface protein A in children. Journal of Infectious Diseases. 2001;183(6):887-96.

51.   Simell B, Kilpi TM, Kayhty H. Pneumococcal carriage and otitis media induce salivary antibodies to pneumococcal capsular polysaccharides in children. The Journal of infectious diseases. 2002;186(8):1106-14. doi: 10.1086/344235. PubMed PMID: 12355361.

52.   Trzcinski K, Thompson CM, Srivastava A, Basset A, Malley R, Lipsitch M. Protection against nasopharyngeal colonization by Streptococcus pneumoniae is mediated by antigen-specific CD4+ T cells. Infection and immunity. 2008;76(6):2678-84. doi: 10.1128/IAI.00141-08. PubMed PMID: 18391006; PubMed Central PMCID: PMCPMC2423086.

53.   Giefing C, Meinke AL, Hanner M, Henics T, Bui MD, Gelbmann D, et al. Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. J Exp Med. 2008;205(1):117-31. doi: 10.1084/jem.20071168. PubMed PMID: 18166586; PubMed Central PMCID: PMCPMC2234372.

54.   Malley R, Srivastava A, Lipsitch M, Thompson CM, Watkins C, Tzianabos A, et al. Antibody-independent, interleukin-17A-mediated, cross-serotype immunity to pneumococci in mice immunized intranasally with the cell wall polysaccharide. Infection and immunity.

2006;74(4):2187-95. doi: 10.1128/IAI.74.4.2187-2195.2006. PubMed PMID: 16552049; PubMed Central PMCID: PMCPMC1418935.

55.   Lu YJ, Gross J, Bogaert D, Finn A, Bagrade L, Zhang Q, et al. Interleukin-17A mediates acquired immunity to pneumococcal colonization. PLoS pathogens. 2008;4(9):e1000159. doi: 10.1371/journal.ppat.1000159. PubMed PMID: 18802458; PubMed Central PMCID: PMCPMC2528945.

56.   Glennie SJ, Banda D, Gould K, Hinds J, Kamngona A, Everett DDB, et al. Defective pneumococcal-specific Th1 responses in HIV-infected adults precedes a loss of control of pneumococcal colonization. Clinical Infectious Diseases. 2013.

57.   Goldblatt D, Mahein H, Andrews N, Ashton L, Virta C, Melegaro A, et al. Antibody responses to nasopharyngeal carriage of Streptococcus pneumoniae in adults: a longitudinal household study. Journal of Infectious Diseases. 2005;192(3):387-93.

58.   Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, Lipsitch M. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. Journal of Infectious Diseases. 2008;197(11):1511-8.

59.   Hill PC, Cheung Y, Akisanya A, Sankareh K, Lahai G, Greenwood BM, et al. Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian infants: a longitudinal study. Clinical Infectious Diseases. 2008;46(6):807-14.

60.   The Green Book. Chapter 25: Pneumococcal. In: England PH, editor. 2013.

61.   Russell FM, Carapetis JR, Satzke C, Tikoduadua L, Waqatakirewa L, Chandra R, et al. Pneumococcal nasopharyngeal carriage following reduced doses of a 7-valent pneumococcal conjugate vaccine and a 23-valent pneumococcal polysaccharide vaccine booster. Clinical and vaccine immunology : CVI. 2010;17(12):1970-6. doi: 10.1128/CVI.00117-10. PubMed PMID: 20943882; PubMed Central PMCID: PMC3008188.

62.   Madhi SA, Nunes MC. The potential impact of pneumococcal conjugate vaccine in Africa: Considerations and early lessons learned from the South African experience. Human vaccines & immunotherapeutics. 2016;12(2):314-25. doi: 10.1080/21645515.2015.1084450. PubMed PMID: 26317537; PubMed Central PMCID: PMCPMC5049711.

63.   Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. The Pediatric infectious disease journal. 2015;34(8):858-64. doi: 10.1097/INF.0000000000000717. PubMed PMID: 26075814.

64.   Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL. The fundamental link between pneumococcal carriage and disease. Expert review of vaccines. 2012;11(7):841-55.

65.   Pichichero ME. Protein carriers of conjugate vaccines: characteristics, development, and clinical trials. Human vaccines & immunotherapeutics. 2013;9(12):2505-23. doi: 10.4161/hv.26109. PubMed PMID: 23955057; PubMed Central PMCID: PMCPMC4162048.

66.   Clarke ET, Williams NA, Findlow J, Borrow R, Heyderman RS, Finn A. Polysaccharide-specific memory B cells generated by conjugate vaccines in humans conform to the CD27+IgG+ isotype-switched memory B Cell phenotype and require contact-dependent signals from bystander T cells activated by bacterial proteins to differentiate into plasma cells. J Immunol. 2013;191(12):6071-83. doi: 10.4049/jimmunol.1203254. PubMed PMID: 24227777.

67.   Davis SM, Deloria-Knoll M, Kassa HT, O'Brien KL. Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: Review of evidence on indirect effects. Vaccine. 2013;32(1):133-45. doi: 10.1016/j.vaccine.2013.05.005. PubMed PMID: 23684824.

68.   Flasche S, Le Polain de Waroux O, O'Brien KL, Edmunds WJ. The serotype distribution among healthy carriers before vaccination is essential for predicting the impact of pneumococcal conjugate vaccine on invasive disease. PLoS computational biology. 2015;11(4):e1004173. doi: 10.1371/journal.pcbi.1004173. PubMed PMID: 25879748; PubMed Central PMCID: PMCPMC4400047.

69.   Frey SE, Lottenbach KR, Hill H, Blevins TP, Yu Y, Zhang Y, et al. A Phase I, dose-escalation trial in adults of three recombinant attenuated Salmonella Typhi vaccine vectors producing Streptococcus pneumoniae surface protein antigen PspA. Vaccine. 2013;31(42):4874-80. doi: 10.1016/j.vaccine.2013.07.049. PubMed PMID: 23916987.

70.   Briles DE, Miyaji E, Fukuyama Y, Ferreira DM, Fujihashi K. Elicitation of mucosal immunity by proteins of Streptococcus pneumoniae. (Advances in Oto-Rhino-Laryngology, Volume 72). Recent advances in tonsils and mucosal barriers of the upper airways; 2011 :25-27 11 ref.

71.   Brooks WA, Chang LJ, Sheng X, Hopfer R, Team PPRS. Safety and immunogenicity of a trivalent recombinant PcpA, PhtD, and PlyD1 pneumococcal protein vaccine in adults, toddlers, and infants: A phase I randomized controlled study. Vaccine. 2015;33(36):4610-7. doi: 10.1016/j.vaccine.2015.06.078. PubMed PMID: 26143615.

72.   Zhang F, Lu YJ, Malley R. Multiple antigen-presenting system (MAPS) to induce comprehensive B- and T-cell immunity. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(33):13564-9. doi: 10.1073/pnas.1307228110. PubMed PMID: 23898212; PubMed Central PMCID: PMCPMC3746895.

73.   Alderson MR. Status of research and development of pediatric vaccines for Streptococcus pneumoniae. Vaccine. 2016;34(26):2959-61. doi: 10.1016/j.vaccine.2016.03.107. PubMed PMID: 27083428; PubMed Central PMCID: PMCPMC4906266.

74.   Johnson HL, Deloria-Knoll M, Levine OS, Stoszek SK, Freimanis Hance L, Reithinger R, et al. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. PLoS medicine. 2010;7(10). doi: 10.1371/journal.pmed.1000348. PubMed PMID: 20957191; PubMed Central PMCID: PMC2950132.

75.   Global Vaccine Introduction Status. Vaccine Information Management Sys-tem (VIMS). Accelerated Vaccine Initiative March, 2013 report. . Johns HopkinsBloomberg School of Public Health. : 2013.

76.   GAVI. Advance Market Commitments for Pneumocococcal Vaccines. Annual Report Geneva: GAVI   Alliance;   2012   [17/12/2012].   Available   from: http://www.gavialliance.org/funding/pneumococcal-amc/.

77.   Usuf E, Bottomley C, Adegbola RA, Hall A. Pneumococcal carriage in sub-Saharan Africa - a systematic review. PloS one. 2014;9(1).

78.   O'Brien KL, Millar EV, Zell ER, Bronsdon M, Weatherholtz R, Reid R, et al. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. The Journal of infectious diseases. 2007;196(8):1211-20. Epub 2007/10/24. doi: 10.1086/521833. PubMed PMID: 17955440.

79.   Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, Reasonover A, et al. Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. JInfectDis. 2006;193(11):1487-94. doi: JID35666 [pii];10.1086/503805 [doi].

80.   Mackenzie GA, Plumb ID, Sambou S, Saha D, Uchendu U, Akinsola B, et al. Monitoring the introduction of pneumococcal conjugate vaccines into West Africa: design and implementation of a population-based surveillance system. PLoS medicine. 2012;9(1):e1001161. doi: 10.1371/journal.pmed.1001161. PubMed PMID: 22272192; PubMed Central PMCID: PMC3260317.

81.   Flasche S, Hoek AJv, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoS medicine. 2011;8(4).

82.   Weinberger DM, Malley R, Lipsitch M. Serotype replacement in disease after pneumococcal   vaccination.   Lancet.   2011;378(9807):1962-73.   Epub   2011/04/16.   doi:

10.1016/S0140-6736(10)62225-8. PubMed PMID: 21492929; PubMed Central PMCID: PMC3256741.

83.   Feikin DR, Kagucia EW, Loo JD, Link-Gelles R, Puhan MA, Cherian T, et al. Serotype-specific changes in Invasive Pneumococcal Disease after Pneumocococal Conjugate Vaccine Introduction: A Pooled Analysis of Multiple Surveillance Sites. PLoS medicine. 2013;10(9): e1001517.

84.   Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. Lancet Global Health. 2014;2(7):e397-e405.

85.   Roca A, Dione MM, Bojang A, Townend J, Egere U, Darboe O, et al. Nasopharyngeal carriage of pneumococci four years after community-wide vaccination with PCV-7 in the Gambia: long-term evaluation of a cluster randomized trial. PloS one. 2013;8(9).

86.   Rodgers GL, Klugman KP. Surveillance of the impact of pneumococcal conjugate vaccines in developing countries. Human vaccines & immunotherapeutics. 2016;12(2):417-20. doi: 10.1080/21645515.2015.1057671. PubMed PMID: 26309055; PubMed Central PMCID: PMCPMC5049713.

87.   Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a   cross-sectional   study.   PLoSMed.   2011;8(4):e1001017.   doi: 10.1371/journal.pmed.1001017 [doi].

88.   Huang SS, Platt R, Rifas-Shiman SL, Pelton SI, Goldmann D, Finkelstein JA. Post-PCV7 changes in colonizing pneumococcal serotypes in 16 Massachusetts communities, 2001 and 2004. Pediatrics. 2005;116(3):e408-13. doi: 10.1542/peds.2004-2338. PubMed PMID: 16140686.

89.   Spijkerman J, van Gils EJ, Veenhoven RH, Hak E, Yzerman EP, van der Ende A, et al. Carriage of Streptococcus pneumoniae 3 years after start of vaccination program, the Netherlands. EmergInfectDis. 2011;17(4):584-91. doi: 10.3201/eid1704.101115 [doi].

90.   Lipsitch M. Vaccination against colonizing bacteria with multiple serotypes. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(12):6571-6. PubMed PMID: 9177259; PubMed Central PMCID: PMC21091.

91.   Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. JInfectDis. 2010;201(1):32-41. doi: 10.1086/648593 [doi].

92.   Miller E, Andrews NJ, Waight PA, Slack MP, George RC. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. The Lancet infectious diseases. 2011;11(10):760-8. doi: 10.1016/S1473-3099(11)70090-1. PubMed PMID: 21621466.

93.   Isaacman DJ, McIntosh ED, Reinert RR. Burden of invasive pneumococcal disease and serotype distribution among Streptococcus pneumoniae isolates in young children in Europe: impact of the 7-valent pneumococcal conjugate vaccine and considerations for future conjugate vaccines. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2010;14(3):e197-209. Epub 2009/08/25. doi: 10.1016/j.ijid.2009.05.010. PubMed PMID: 19700359.

94.   Mackenzie GA, Hill PC, Jeffries DJ, Hossain I, Uchendu U, Ameh D, et al. Effect of the introduction of pneumococcal conjugate vaccination on invasive pneumococcal disease in The Gambia: a population-based surveillance study. The Lancet infectious diseases. 2016;16(6):703-11. doi: 10.1016/S1473-3099(16)00054-2. PubMed PMID: 26897105; PubMed Central PMCID: PMCPMC4909992.

95.   Feikin DR, Kagucia EW, Loo JD, Link-Gelles R, Puhan MA, Cherian T, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a   pooled   analysis   of   multiple   surveillance   sites.   PLoS   medicine.

2013;10(9):e1001517. doi: 10.1371/journal.pmed.1001517. PubMed PMID: 24086113; PubMed Central PMCID: PMC3782411.

96.   Loo JD, Conklin L, Fleming-Dutra KE, Deloria Knoll M, Park DE, Kirk J, et al. Systematic review of the effect of pneumococcal conjugate vaccine dosing schedules on prevention of pneumonia. Pediatr Infect Dis J. 2014;33 Suppl 2:S140-51. doi: 10.1097/INF.0000000000000082. PubMed PMID: 24336056; PubMed Central PMCID: PMC3944478.

97.   Hammitt LL, Ojal J, Bashraheil M, Morpeth SC, Karani A, Habib A, et al. Immunogenicity, impact on carriage and reactogenicity of 10-Valent pneumococcal non-typeable Haemophilus influenzae protein D conjugate vaccine in Kenyan children aged 1-4 years: a randomized controlled trial. PloS one. 2014;9(1).

98.   Roca A, Bojang A, Bottomley C, Gladstone RA, Adetifa JU, Egere U, et al. Effect on nasopharyngeal pneumococcal carriage of replacing PCV7 with PCV13 in the Expanded Programme of Immunization in The Gambia. Vaccine. 2015;33(51):7144-51. doi: 10.1016/j.vaccine.2015.11.012. PubMed PMID: 26592141.

99.   Woolfson A, Huebner R, Wasas A, Chola S, Godfrey-Faussett P, Klugman K. Nasopharyngeal carriage of community-acquired, antibiotic-resistant Streptococcus pneumoniae in a Zambian paediatric population. Bulletin of the World Health Organization. 1997;75(5):453-62. PubMed PMID: 9447779; PubMed Central PMCID: PMC2487017 years of age in sub-Saharan Africa. This study investigated the nasopharyngeal carriage rate of Streptococcus pneumoniae in a Zambian pediatric population and the prevalence of antibiotic resistance. Enrolled were 260 children under 6 years of age (mean age, 20 months) treated at the University Teaching Hospital in Lusaka, Zambia, in 1994. S. pneumoniae was isolated from the nasopharynx of 187 children (71.9%). The odds of carrying pneumococci were twice as high among children under 2 years of age (76.2%) than older children (59.7%). Overall, 83 (65.9%) of the 126 isolates available for antibiotic resistance profiles were sensitive to the drugs. Resistance to tetracycline, penicillin, sulfamethoxazole plus trimethoprim, and chloramphenicol was found in 23.0%, 14.3%, 12.7%, and 3.9%, respectively, of the isolates. The highest level of resistance was recorded in all isolates resistant to tetracycline. All but one of the multidrug-resistant isolates were serotype 14. Children under 6 months old were least likely to carry antibiotic-resistant organisms. In an anonymous questionnaire completed by 160 mothers, 38% reported they obtained antibiotics without a prescription and 49.4% acknowledged feeling dissatisfied when not given antibiotics to treat their sick child. Ongoing surveillance is recommended in Zambia to ensure that recommended treatment regimens keep pace with trends in antibiotic resistance.

100.   Publication WHO. Pneumococcal vaccines WHO position paper - 2012 - recommendations. Vaccine. 2012;30(32):4717-8. doi: 10.1016/j.vaccine.2012.04.093. PubMed PMID: 22621828.

101.   Scott P, Rutjes AW, Bermetz L, Robert N, Scott S, Lourenco T, et al. Comparing pneumococcal conjugate vaccine schedules based on 3 and 2 primary doses: systematic review and meta-analysis. Vaccine. 2011;29(52):9711-21. doi: 10.1016/j.vaccine.2011.07.042. PubMed PMID: 21821080.

102.   Gavi, The Vaccine Alliance, ,. Advance Market Commitment for Pneumococcal Vaccines. Annual Report 1 April 2014 - 31 March 2015. 2015.

103.   Whitney CG, Goldblatt D, O'Brien KL. Dosing schedules for pneumococcal conjugate vaccine: considerations for policy makers. The Pediatric infectious disease journal. 2014;33 Suppl 2:S172-81. doi: 10.1097/INF.0000000000000076. PubMed PMID: 24336059.

104.   Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. The Lancet Global health. 2014;2(7):e397-405. doi: 10.1016/S2214-109X(14)70224-4. PubMed PMID: 25103393.

105.   Cameron C, Pebody R. Introduction of pneumococcal conjugate vaccine to the UK childhood immunisation programme, and changes to the meningitis C and Hib schedules. Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin. 2006;11(3):E060302 4. PubMed PMID: 16804223.

106.   Picon T, Alonso L, Garcia Gabarrot G, Speranza N, Casas M, Arrieta F, et al. Effectiveness of the 7-valent pneumococcal conjugate vaccine against vaccine-type invasive disease among children in Uruguay: an evaluation using existing data. Vaccine. 2013;31 Suppl 3:C109-13. doi: 10.1016/j.vaccine.2013.01.059. PubMed PMID: 23777683.

107.   WHO. Draft Vaccine Action Plan. Report by the Secretariat. Sixty-fifth World Health Assembly   11 May 2012.

108.   Heesterbeek H, Anderson RM, Andreasen V, Bansal S, De Angelis D, Dye C, et al. Modeling infectious disease dynamics in the complex landscape of global health. Science. 2015;347(6227):aaa4339. doi: 10.1126/science.aaa4339. PubMed PMID: 25766240; PubMed Central PMCID: PMCPMC4445966.

109.   Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine. 2008;5(3):e74. doi: 10.1371/journal.pmed.0050074. PubMed PMID: 18366252; PubMed Central PMCID: PMC2270306.

110.   Edmunds WJ, O'Callaghan CJ, Nokes DJ. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proceedings Biological sciences / The Royal Society. 1997;264(1384):949-57. doi: 10.1098/rspb.1997.0131. PubMed PMID: 9263464; PubMed Central PMCID: PMC1688546.

111.   Goldblatt D, Ramakrishnan M, O'Brien K. Using the impact of pneumococcal vaccines on nasopharyngeal carriage to aid licensing and vaccine implementation; a PneumoCarr meeting report March 27-28, 2012, Geneva. Vaccine. 2013;32(1):146-52. doi: 10.1016/j.vaccine.2013.06.040. PubMed PMID: 23933374.

112.   McLean AR. Vaccination, evolution and changes in the efficacy of vaccines: a theoretical framework. Proceedings Biological sciences / The Royal Society. 1995;261(1362):389-93. doi: 10.1098/rspb.1995.0164. PubMed PMID: 8587880.

113.   Nowak MA, May RM. Superinfection and the evolution of parasite virulence. Proceedings Biological sciences / The Royal Society. 1994;255(1342):81-9. doi: 10.1098/rspb.1994.0012. PubMed PMID: 8153140.

114.   May RM, Nowak MA. Superinfection, metapopulation dynamics, and the evolution of diversity. Journal of theoretical biology. 1994;170(1):95-114. doi: 10.1006/jtbi.1994.1171. PubMed PMID: 7967636.

115.   Choi YH, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, et al. 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PLoSOne. 2011;6(10):e26190. doi: 10.1371/journal.pone.0026190 [doi];PONE-D-11-09966 [pii].

116.   Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, Gay NJ. Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMCInfectDis. 2010;10:90. doi: 1471-2334-10-90 [pii];10.1186/1471-2334-10-90 [doi].

117.   Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. PLoSOne. 2012;7(7):e39927. doi: 10.1371/journal.pone.0039927 [doi];PONE-D-12-11977 [pii].

118.   De Cao E, Melegaro A, Klok R, Postma M. Optimising assessments of the epidemiological impact in The Netherlands of paediatric immunisation with 13-valent pneumococcal conjugate vaccine using dynamic transmission modelling. PloS one. 2014;9(4):e89415. doi: 10.1371/journal.pone.0089415. PubMed PMID: 24694656; PubMed Central PMCID: PMCPMC3973563.

119.   Bottomley C, Roca A, Hill PC, Greenwood B, Isham V. A mathematical model of serotype replacement in pneumococcal carriage following vaccination. Journal of the Royal Society, Interface / the Royal Society. 2013.

120.   Nurhonen M, Cheng AC, Auranen K. Pneumococcal transmission and disease in silico: a microsimulation model of the indirect effects of vaccination. PloS one. 2013;8(2).

121.   Van Effelterre T, Moore MR, Fierens F, Whitney CG, White L, Pelton SI, et al. A dynamic model of pneumococcal infection in the United States: implications for prevention through vaccination. Vaccine. 2010;28(21):3650-60. doi: 10.1016/j.vaccine.2010.03.030. PubMed PMID: 20359560.

122.   Wu DB, Chang CJ, Huang YC, Wen YW, Wu CL, Fann CS. Cost-effectiveness analysis of pneumococcal conjugate vaccine in Taiwan: a transmission dynamic modeling approach. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2012;15(1 Suppl):S15-9. doi: 10.1016/j.jval.2011.11.013. PubMed PMID: 22265061.

123.   Zhang Y, Auranen K, Eichner M. The influence of competition and vaccination on the coexistence of two pneumococcal serotypes. Epidemiology and infection. 2004;132(6):1073-81. PubMed PMID: 15635964; PubMed Central PMCID: PMC2870198.

124.   Huang SS, Finkelstein JA, Lipsitch M. Modeling community- and individual-level effects of child-care center attendance on pneumococcal carriage. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2005;40(9):1215-22. doi: 10.1086/428580. PubMed PMID: 15825020.

125.   Karlsson D, Jansson A, Normark BH, Nilsson P. An individual-based network model to evaluate interventions for controlling pneumococcal transmission. BMC infectious diseases. 2008;8:83. doi: 10.1186/1471-2334-8-83. PubMed PMID: 18559109; PubMed Central PMCID: PMC2442080.

126.   Erasto P, Hoti F, Auranen K. Modeling transmission of multitype infectious agents: application to carriage of Streptococcus pneumoniae. Statistics in medicine. 2012;31(14):1450-63. doi: 10.1002/sim.4487. PubMed PMID: 22354452.

127.   Lipsitch M, Abdullahi O, D'Amour A, Xie W, Weinberger DM, Tchetgen Tchetgen E, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. Epidemiology. 2012;23(4):510-9. Epub 2012/03/24. doi: 10.1097/EDE.0b013e31824f2f32. PubMed PMID: 22441543.

128.   Melegaro A, Gay NJ, Medley GF. Estimating the transmission parameters of pneumococcal carriage in households. Epidemiology and infection. 2004;132(3):433-41.

129.   Flasche S, Edmunds J, Millar E, Goldblatt D, Robertson C, Choi JY. The impact of specific and non-specific immunity on the ecology of Streptoccus pneumoniae and the implications for vaccination. Proc R Soc B. 2013;(20131939).

130.   De Wals P, Black S, Borrow R, Pearce D. Modeling the impact of a new vaccine on pneumococcal and nontypable Haemophilus influenzae diseases: a new simulation model. Clinical therapeutics. 2009;31(10):2152-69. doi: 10.1016/j.clinthera.2009.10.014. PubMed PMID: 19922887.

2. Research paper 1: Age-dependent prevalence of nasopharyngeal carriage of Streptococcus pneumoniae before conjugate vaccine introduction: a prediction model based on a meta-analysis

Olivier le Polain de Waroux[1], Stefan Flasche[1], David Prieto-Merino[2], W. John Edmunds[1]

1. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK
2. Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Olivier le Polain de Waroux
Principal Supervisor	W John Edmunds
Thesis Title	Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	PLOS One		
When was the work published?	2014		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	NA
Please list the paper's authors in the intended authorship order:	NA
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate conceived the project, designed the study, explored statistical methods to address the question. Input was provided by co-authors in the design and statistical approach. The candidate compiled the data, performed the analysis and wrote the first draft of the manuscript. All co-authors

	reviewed the manuscript and approved its final version

Student Signature: _____ **Date:** 18 November 2016

Supervisor Signature: _____ **Date:** 18 November 2016

Abstract

Introduction

Data on the prevalence of nasopharyngeal carriage of *S.pneumoniae* in all age groups are important to help predict the impact of introducing pneumococcal conjugate vaccines (PCV) into routine infant immunization, given the important indirect effect of the vaccine. Yet most carriage studies are limited to children under five years of age. We here explore the association between carriage prevalence and serotype distribution in children aged ≥5 years and in adults compared to children.

Methods

We conducted a systematic review of studies providing carriage estimates across age groups in healthy populations not previously exposed to PCV, using MEDLINE and Embase. We used Bayesian linear meta-regression models to predict the overall carriage prevalence as well as the prevalence and distribution of vaccine and nonvaccine type (VT and NVT) serotypes in older age groups as a function of that in <5y olds.

Results

Twenty-nine studies compromising of 20,391 individuals were included in the analysis. In all studies nasopharyngeal carriage decreased with increasing age. We found a strong positive linear association between the carriage prevalence in pre-school childen (<5y) and both that in school aged children (5 – 17y olds) and in adults. The proportion of VT serotypes isolated from carriers was consistently lower in older age groups and on average about 73% that of children <5y among 5-17y olds and adults respectively. We provide a prediction model to infer the carriage prevalence and serotype distribution in 5-17y olds and adults as a function of that in children <5 years of age.

Conclusion

Such predictions are helpful for assessing the potential population-wide effects of vaccination programmes, e.g. via transmission models, and thus assist in the design of future pneumococcal conjugate vaccination strategies.

2.1.   Introduction

Colonization of the nasopharynx by *Streptococcus pneumoniae* is the reservoir for *S.pneumoniae* transmission and a prerequisite for pneumococcal disease [1]. Pneumococcal conjugate vaccines (PCV) reduce nasopharyngeal carriage of serotypes included in the vaccine by conferring capsular-specific immunity. Experience from countries where conjugate vaccines have been introduced has shown rapid and sustained carriage reduction of vaccine serotypes (VT) following vaccination. Those trends have been observed not only among vaccinated children but more widely across all age groups through a strong herd immunity effect [2,3]. Despite evidence of almost complete serotype replacement in many settings, whereby non-vaccine serotypes (NVT) colonise the space left vacant by vaccine type (VT) serotypes [4], pneumococcal conjugate vaccination programmes have led to a substantial reduction in severe disease due to the lower propensity of replacing serotypes to cause disease [4,5].

Ten- and thirteen-valent pneumococcal conjugate vaccines (PCV10 and PCV13) are now being introduced into the routine immunization programmes of many developing countries (www.jhsph.edu/ivac/vims.html), where their impact is expected to be high, given the disproportionate burden of pneumococcal disease in such settings [6].

Estimates of the potential impact of routinely introducing pneumococcal conjugate vaccines (PCVs), however, crucially depend on the nasopharyngeal carriage prevalence in the population before the introduction of PCV, the distribution of serotypes (VT and NVT) within the population, including among older children, adults and the elderly, and the propensity of replacing serotypes to cause disease across age groups.

Most carriage surveys are limited to children under five years of age, in whom the disease burden is high and for whom sample size requirements for precision are reasonable given the high carriage prevalence. As a result, nasopharyngeal carriage estimates in other age groups are scarce. However, such estimates are important to help predict the overall population impact of vaccination programmes as well as the specific impact among unvaccinated age groups.

Routine infant PCV vaccination has been found to also impact substantially on the elderly in whom the likelihood to develop severe pneumococcal disease as a result of carriage is high [7,8] and who present the highest overall burden of pneumococcal associated disease in developed countries [9]. Hence this age group has also played an important role in the cost effectiveness considerations of pneumococcal conjugate vaccination [10,11]. With the proportion of population of >60 years old growing at its fastest pace ever [12], the indirect effect of PCV vaccination programmes may become increasingly important, including in developing countries.

The overall aim of this study was to explore a possible correlation between the prevalence and distribution of *S.pneumoniae* serotypes carried in the nasopharynx of children <5y and that in older children (5-17y olds) and adults (≥18y olds), based on nasopharyngeal carriage surveys, and further establish predictors for carriage prevalence and serotype distribution in adults and older children as a function of the carriage prevalence and serotype distribution in the nasopharynx of children under five years of age.

2.2. Methods

2.2.1. Search strategy

We conducted a systematic review to identify articles reporting nasopharyngeal carriage prevalence estimates for different age strata. We used MEDLINE and Embase electronic databases to retrieve articles published between the date of the earliest articles compiled on MEDLINE (1946) or Embase (1947) and 23rd August 2013 (i.e. week 35), and used the following combination of search terms: '(pneumonia OR pneumoniae OR pneumococcal OR pneumococcus) AND (carriage OR colonization OR colonisation)' in the title or the keywords or the abstract. No language restriction was applied.

There is no registered protocol for this systematic review.

Our systematic review and meta-analysis was conducted in accordance with the PRISMA checklist (http://www.prisma-statement.org/statement.htm, see Supporting Information) and MOOSE guidelines [13], which compile guidelines for the reporting of meta-analysis of observational studies.

2.2.2. Eligibility criteria

We included studies based on seven main eligibility criteria.

Articles were considered for inclusion if they provided (i) pneumococcal nasopharyngeal carriage prevalence estimates (ii) in a population not previously exposed to PCV, with (iii) nasopharyngeal sampling and transport procedures as well as *S.pneumoniae* culture based on WHO guidelines [14], (iv) where the study was not restricted to specific serotypes or to *S.pneumoniae* with specific patterns of antibiotic sensitivity. Studies were further considered for inclusion if they provided carriage estimates in young children, as well as in older age groups and (vi) in individuals not suffering from any acute respiratory infection or any confirmed pneumococcal disease, and (vii) were not based on particular at risk population groups such as HIV positive individuals.

No design restriction was applied.

2.2.3. Data extraction

The articles were screened and reviewed with inclusion criteria appraised in the order described above. When primary data published in a study were also used in subsequent studies, we screened the latter too to find any data that may not have been published in the original paper. For each study meeting the aforementioned inclusion criteria we calculated the prevalence of nasopharyngeal carriage by age group, as well as the prevalence of VT and NVT by age group when provided, for PCV7 and/or PCV10 and/or PCV13, depending on available data. In most

studies the group of NVT comprised of NVT serotypes as well as non-typeable (NT) serotypes, while in a few studies estimates for NT by age group were provided separately and were therefore not included in the group of NVT serotypes.

In some studies, estimates were provided by age or for smaller age bands, and such estimates were therefore pooled to obtain estimates for the main four age groups considered.

In longitudinal studies where multiple nasopharyngeal swabs were taken for each individual, the number of individuals tested positive was approximated by the age-specific average number of positive swabs over the study period.

In a few studies the actual number of carriers had to be estimated based on reported prevalence estimates and the number of study participants in each age group.

More details on how the data were extracted from the different studies can be found in the Supporting Information.

2.2.4. Analysis

We considered the following age groups: <1y (infants); <5y (pre-school children including infants); 5-17y (school aged children) and ≥18y (adults). Because age groups were not standardised between studies, the category of <5y olds included studies reporting estimates in <4y olds as well as studies reporting estimates in <6y olds. The category of school-aged children included any age group from between 4 to 6 years up to any age between 10 years and 19 years, and we considered the prevalence among adults to be that in individuals aged at least 15 years and above.

We explored the association between the carriage prevalence and VT or NVT distribution in older age groups and that in young children using Bayesian linear meta-regression analysis. The use of Bayesian over a frequentist approach was motivated by the natural way in which each study's contribution to the meta-regression is weighted in a Bayesian approach, and also because Bayesian linear regression is the recommended tool by the Cochrane Collaboration to

account for uncertainty around both the outcome and the exposure variables in a meta-regression [15].

2.2.5. Age-dependent overall carriage prevalence and carriage prediction

We obtained prediction intervals for the carriage prevalence in adults and in 5-17y olds as a function of that in either <5y olds or <1y olds using a Bayesian linear meta- regression model. For the i studies included in each analysis,

$$P_{Ai} = \beta_0 + \beta_1 * P_{Ci} + \varepsilon_i$$

, with P_{Ai} = prevalence in either adults or 5-17y olds in study i, P_{Ci} =prevalence in either <5y olds or <1y olds in study i and ε_i= random error in study $\varepsilon_i \sim N(0, \sigma^2)$

The true prevalence P_{Ai} and P_{Ci} are unknown, however the observed number of carriers in each study (X_{Ai} and X_{Ci}) follows a binomial distribution. Hence, based on these and on the sample sizes (N_{Ai} and N_{Ci}) it follows that $X_{Ai} \sim$ Binomial(P_{Ai}, N_{Ai}) and $X_{Ci} \sim$ Binomial(P_{Ci}, N_{Ci})

We provided prior distributions for the parameters of interest β_0, β_1 and σ and assigned uniform uninformative priors to β_0 (unif (-1,1)) and β_1 (unif (-5,5)) and to σ (unif (0,0.4)).

The posterior distributions were obtained through a Markov Chain Monte Carlo (MCMC) Gibbs sampling algorithm, with 100,000 iterations of 2 chains running in parallel, after a burn-in of 5,000 iterations. We retained one in five iterations in the posterior sample to limit autocorrelation. Convergence of the chains was examined visually. We obtained the 95% posterior credible intervals (CrI) for the model regression line as well as the posterior credible intervals for the prediction of individual studies by including the prediction error (which we will refer to as the "prediction interval" in the results section).

We further explored the possible association of socio-demographic and geographic characteristics of the populations studied through Bayesian multivariable meta-regression. The explanatory variables considered for meta-regression included the proportion of children <5y of age and the % <15y of age in the country or area in which the study took place based on figures obtained from UN Population data (http://esa.un.org/wpp/), categories of national income level in the country or area in which the study took place, based on estimates from the World Bank (http://data.worldbank.org/), and broad geographical areas or continents in which studies took place. Variables were added one after another in the model and were retained if the 95% posterior probability interval for their coefficient excluded 0.

2.2.6. Age-dependent distribution of VT serotypes

We then explored (i) the distribution of the group of VT serotypes among carried serotypes in children <5y and in older age groups, and the association between such distributions, as well as (ii) the relationship between the prevalence of VT carriage and NVT carriage in adults and in 5-17y olds compared to children <5 years of age

For each study providing serotype-specific information we calculated (i) the proportion of VT serotypes (for PCV7 or PCV10 or PCV13) among carriers in children and older age group, and their 95% confidence interval (CI) and (ii) the carriage prevalence of such groups of serotypes, by age category and 95% CI.

We further explored the relationship between (i) the proportion of VT carriers and (ii) the VT and NVT carriage prevalence across age groups using Bayesian linear meta-regression. We used the same uniform priors for $\beta 0$, $\beta 1$ and σ^2 than in the analysis of overall carriage

prevalence, as well as the same analytical strategy to obtain posterior estimates.

The code used for the Bayesian linear meta-regression can be found in the Supporting Information and is fully annotated. In addition, the code also provides an opportunity for

readers to obtain posterior distributions of the carriage prevalence in 5-17y olds and adults based on study-specific estimates of nasopharyngeal carriage in <5y olds, making it possible to use this as a carriage prediction tool based on specific data of carriage in <5y olds.

Analyses were performed using R software and the JAGS package in R (http://mcmc-jags.sourceforge.net/).

2.3. Results

A total of 8,886 citations were found, which amounted to 4,648 citations after duplicates were excluded. Of those, 376 original studies provided pre-PCV nasopharyngeal carriage estimates in healthy individuals. A flowchart of the number of studies screened and reasons for exclusion is displayed in Figure 1.

A total of 29 studies were included in the meta-analysis, comprising a total of 20,391 individuals, including 7931 children <5y, 3936 school aged children and 8524 adults.

2.3.1. Age dependent prevalence of nasopharyngeal carriage

Seven studies were conducted in Africa, 6 in Asia, 6 in the Middle-East, 5 in Europe, 2 in the North America (Alaska, USA) and 3 in South America. Table 1 displays the main characteristics of the studies included in the final analysis, as well as the estimates of carriage prevalence by age group.

We found a strong positive correlation between carriage prevalence in younger age groups and that in older age groups. Figure 2 displays a scatter plot of the study specific estimates for the prevalence in adults and in 5-17y olds as a function of that in either <5y olds or <1y olds.

Figure 1: Flow chart of the study selection process.

```
┌─────────────────────────────────────────────┐
│ A total of 8,886 citations were identified:   │
│   -   4,765 citations in Embase               │
│   -   4,121 citations in MEDLINE              │
└─────────────────────────────────────────────┘
                    │
                    │      ┌─────────────────────────────────────────────┐
                    ├─────▶│ 4,238 citations excluded after de-duplication │
                    │      └─────────────────────────────────────────────┘
                    ▼
┌─────────────────────────────────────────────┐
│ 4,648 studies screened for possible inclusion │
└─────────────────────────────────────────────┘
                    │
                    │      ┌──────────────────────────────────────────────────────────────┐
                    │      │ 3,992 studies excluded as they provided no carriage prevalence │
                    │      │ estimates in humans                                            │
                    ├─────▶│   -   3,633 excluded based on title only                       │
                    │      │   -   331 excluded based on title + abstract                   │
                    │      │   -   28 excluded after reading full text                      │
                    │      └──────────────────────────────────────────────────────────────┘
                    ▼
┌─────────────────────────────────────────────┐
│ 656 studies provided nasopharyngeal carriage  │
│ prevalence estimates in humans                │
└─────────────────────────────────────────────┘
                    │
                    │      ┌──────────────────────────────────────────────────────────────────────────┐
                    │      │ 280 studies excluded:                                                      │
                    │      │   -   117 studies conducted among ill individuals (90 excluded based on     │
                    │      │       abstract only and 27 based on the full text)                          │
                    │      │   -   73  studies based on secondary data (31 excluded based on abstract     │
                    ├─────▶│       only and 42 after reading full text)                                  │
                    │      │   -   87 primary studies with post-PCV estimates only (22 excluded based on  │
                    │      │       abstract only and 65 after reading full text)                         │
                    │      │   -   3  studies among children <5y with unclear health status of            │
                    │      │       participants or unclear sampling dates (pre- or post-PCV) (2 excluded  │
                    │      │       based on abstract only,1 after reading full text)                     │
                    │      └──────────────────────────────────────────────────────────────────────────┘
                    ▼
┌─────────────────────────────────────────────┐
│ 376 original studies providing pre-PCV NP     │
│ carriage estimates in healthy individuals     │
│ (355 studies providing pre-PCV estimates only │
│ and 22 providing both pre- and post-PCV       │
│ estimates)                                    │
└─────────────────────────────────────────────┘
                    │
                    │      ┌──────────────────────────────────────────────────────────────────────────┐
                    │      │ 327 studies excluded as they provided carriage estimates in specific age    │
                    │      │ groups only                                                                │
                    │      │   -   289 in children <5y only (155 excluded based on abstract only and 134  │
                    ├─────▶│       after reading full text)                                             │
                    │      │   -   10 in 5-17y olds (4 excluded based on abstract only and 6 after        │
                    │      │       reading full text                                                    │
                    │      │   -   27 among adults only (13 excluded based on abstract only and 14 after  │
                    │      │       reading full text)                                                   │
                    │      └──────────────────────────────────────────────────────────────────────────┘
                    ▼
┌─────────────────────────────────────────────┐
│ 50 studies considered for inclusion           │
└─────────────────────────────────────────────┘
                    │
                    │      ┌──────────────────────────────────────────────────────────────────────────┐
                    │      │ 21 studies further excluded (based on full text)                           │
                    │      │   -   3 studies with age group breakdown not compatible with our analysis   │
                    │      │   -   8 studies where age-specific carriage prevalence estimates could not   │
                    │      │       be obtained from the paper or conference abstract                    │
                    │      │   -   1 study with HIV positive adults and healthy children                 │
                    │      │   -   2 studies with laboratory methods differing from the WHO standards     │
                    │      │   -   1 study with adult samples obtained in the post-PCV era from clinical  │
                    ├─────▶│       specimens                                                            │
                    │      │   -   2 studies later found to be based on primary data from already         │
                    │      │       included studies                                                     │
                    │      │   -   1 study with carriage prevalence estimates provided for resistant      │
                    │      │       strains only                                                         │
                    │      │   -   2 studies based on patients attending outpatient or emergency          │
                    │      │       departments, in which age-specific carriage estimates for patients    │
                    │      │       without respiratory illness could not be obtained                    │
                    │      │   -   1 study with estimates in children <6 months old only                 │
                    │      └──────────────────────────────────────────────────────────────────────────┘
                    ▼
┌─────────────────────────────────────────────┐
│            29 studies included                │
└─────────────────────────────────────────────┘
```

The figure also displays the fitted regression line from the Bayesian linear meta-regression model, including the median posterior estimate, the 95% credible interval around the median, as well as the 95% prediction intervals for all four analyses. The model coefficients of all four models can be found in Table 2.

We explored whether the results from the regression model were confounded by the socio-demographic and geographic characteristics of the study population, but found no evidence that the proportion of children, the national income level or the geographic region of the included studies were associated with the outcome in any of the models considered.

Figure 2: Overall carriage prevalence in older age groups against <5y olds: scatter plot and fitted model (<5y (left panels) and <1y (right panels) vs. 5 – 17 y olds (upper panels) and adlults (lower panels))
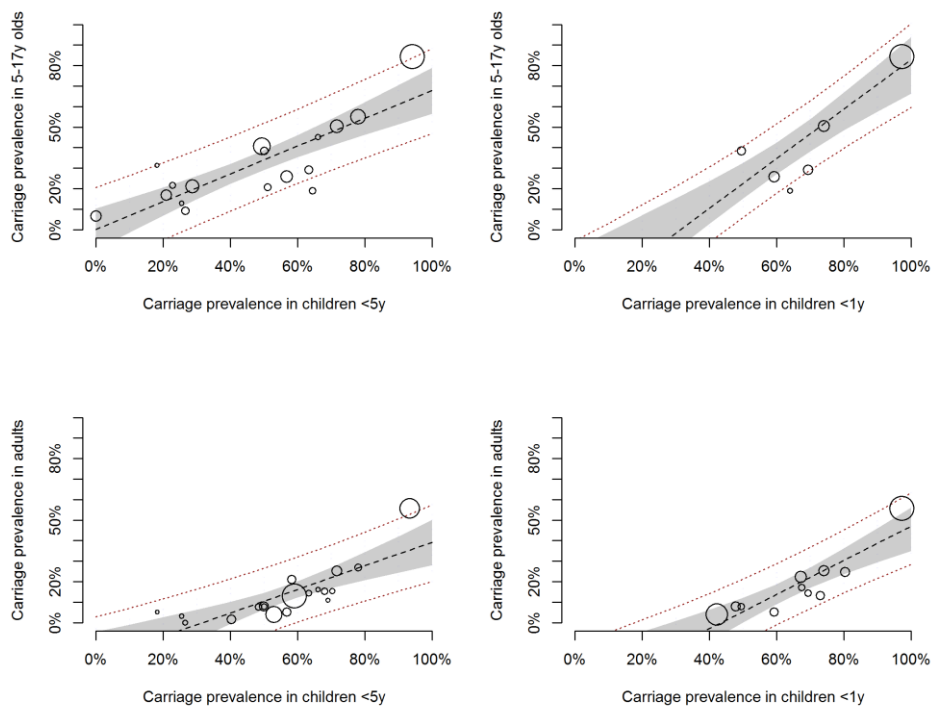
Legend: Each circle corresponds to one study, with the circle size proportional to the study size (i.e. number of individuals contributing to the x and y axis estimates). The lines correspond to the fitted Bayesian linear meta-regression model. The dashed black line shows the median posterior estimate and the grey shaded area the 95% credible interval around the median. The red dotted lines represent the 95% prediction interval.

Table 1: Studies included in the Bayesian linear meta-regression of the overall carriage prevalence

Authors	Country or region	Setting	No carriers/total individuals (%), by age group					Study population characteristics by age group			
			<1y olds	<5y olds	5-17y olds	≥18 years	Age groups	<1y	<5y	5-17y	≥18y
Abdullahi et al. (2008) [17]	Kenya	C	58/98 (59%)	198/349 (57%)	55/213 (55%)	16/302 (5%)	<1y;<5y; 5-19y; ≥20y	RX	RX	RX	RX
Adetifa IM., et al. (2012) [34]	Nigeria	C	143/193(74%)	375/524 (72%)	63/125(25%)	90/356 (25%)	<1y;<5y; 5-14y; ≥15y		RX	RX	RX
Bello Gonzales et al. (2010) [35]	Venezuela	C		58/84 (69%)		7/64 (11%)	<5y; ≥18y		RX		M
Cekmez et al. (2009) [36]	Turkey	H		0/125 (0%)*	25/375 (7%)		<4y ; 4-12y		RH	RH	
Chen et al. (2007) [37]	Taiwan	C		25/94 (27%)	18/196 (9%)	0/137 (0%)	3-4; 5-17y; ≥18y		SC	SC	HC
Dagan et al. (2000) [38]	Israel	C		59/84 (70%)	71/199 (35%)	27/174 (15%)	<6y; 6-15y;≥16y		RX	RX	RX
Darboe MK., et al. (2007) [39]*	Gambia	C	143/196 (73%)			26/196 (13%)	<1y; ≥18y	BC			M
Dhakal et al. (2010) [40]	India	H		18/79 (23%)	26/120 (22%)		<4y ; 4-12y		RH	RH	
Granat et al. (2007) [41]	Bangladesh	C	49/99 (49%)	86/172 (50%)	45/117 (38%)	12/154 (12%)	<1y;<5y;5-18y;≥19y	BC	F	F	F
Greenberg et al. (2004) [42]	Israel	H		147/216 (68%)		33/216(15%)	<5y; ≥18y		RH		M
Hammitt et al. (2006) [3]	Alaska (USA)	C		377/639 (59%)		275/2115 (13%)	<5y; ≥18y		RX		RX
Henriqus Normark et al.(2003) [43]	Sweden	C		246/611 (40%)		2/123 (2%)	1-6y; 19-59y		SC		SC
Hill PC., et al. (2006)[19]	Gambia	C	141/145(97%)	621/666 (93%)	621/735 (84%)	821/1471 (55%)	<1y;<5y; 5-14y; ≥15y	RX	RX	RX	RX
Hussain et al. (2005) [18]**	UK	C		87/180 (48%)	15/71 (21%)	18/237 (8%)	<5y; 5-17y; ≥18y		RX	RX	RX
Inostroza et al.(1998) [44]	Chile	C		10/55 (18%)	5/16 (31%)	2/38 (5%)	<5y; 5-15y; ≥18y		SC	SC	HC
Kaltoft et al. (2008) [45]	Denmark	C		340/584 (58%)		23/109 (21%)	mean 23m (n=123) and mean 52m (n=461); ≥18y		SC		F& SC
Leino et al. (2008) [46]	Finland	C		15/59 (25%)	4/31 (13%)	4/123 (3%)	mean 4y, 8y 35y		SC	F	F& SC
Lloyd-Evans N., et al. (1996) [47]	Gambia	C		323/414 (78%)	188/342 (55%)	18/67 (27%)	<5y;5-18y; ≥19y		RX	F	F
Lo et al. (2003) [48]	Taiwan	H		75/360 (21%)	20/118 (17%)		<6y; 6-14y		RH	RH	
Mueller et al. (2012) [49]	Burkina Faso	C	43/62 (69%)	81/128 (63%)	57/196 (29%)	28/195 (14%)	<1y;<5y; 5-19y; ≥20y	RX	RX	RX	RX
Nunes et al. (2013) [50]	South Africa	C	83/123 (69%)			21/123 (17%)	<16m; ≥18y	BC			M
Parry et al. (2000) [51]	Vietnam	C		192/389 (49%)	212/522 (41%)		<5y; 5-16y		RX	RX	
Regev-Yochay et al. (2004) [16]	Israel	H	38/90 (42%)	214/404 (53%)		52/1300 (4%)	<1y;<7y;≥18y	RH	RH		RH
Regev-Yochay et al. (2012) [52]	Occup. Palest Terr.[2]	C	43/90 (48%)	189/379 (50%)		30/376 (8%)	<1y;<5y; ≥18y	RX	RX		F
Reichler et al. (1992) [53]	USA	C	16/25 (63%)	107/166 (64%)	10/53 (19%)		<18m;<5.5y; 5-10y	SC	SC	SC	
Reis JN. et al. (2008) [54]	Brazil	C		33/50 (66%)	43/95 (45%)	19/117 (16%)	<5y; 5-17y; ≥18y		RX	RX	RX
Sener et al. (1998) [55]	Turkey	C		71/248 (29%)	87/412 (21%)		<6y;6-11y		SC	SC	
Turner et al. (2012) [56]	Thailand[1]	C	188/234 (80%)			57/231 (25%)	<2y; ≥18y	BC			M
van Gils E. et al (2009) [57]	Netherlands	C	214/319 (67%)			67/300 (22%)	12m ; ≥18y	BC			F

H: Health care setting, C: Community setting [1]Maela refugee camp at the Thailand-Myanmar border. [2]Gaza strip. *Data from the study extracted from Darboe et al. [58]. ** Data were extracted from Flasche et al. [20]. RX: Random sample from a cross section survey; RH: random outpatient sample; SC: school cohort (i.e. all children or staff or a particular school/DCC/class); M: Mothers, F: Family members; BC: Birth cohort. Italicised are the samples where adults are parents or staff members looking after the young children included

Specific estimates in older adults were provided in two studies included in the analysis, one in Israel with data in >65y olds [16] and another in Kenya with data in >50y olds [17]. Based on the reported data there was no evidence that the carriage prevalence among older adults differed substantially from that in younger adults (Israel: 3.7% in 18-65y vs. 4.6% in >65y, p=0.573 and Kenya: 5.6% in 20-49y olds vs. 4.6% in ≥50y olds, p=0.719).

Table 2: Model coefficients for each of the Bayesian linear meta-regression models used

Model	Intercept (β_0) median (95% CrI)	Slope (β_1) median (95% CrI)
Overall carriage prevalence		
Adults vs. <5y	-0.18 (-0.31; -0.05)	0.57 (0.35;0.80)
Adults vs. <1y	-0.36 (-0.49; -0.15)	0.83 (0.52; 1.03)
5-17y vs. <5y	-0.00 (-0.10 ; 0.10)	0.68 (0.48 ; 0.86)
5-17y vs. <1y	-0.38 (-0.49 ; -0.09)	1.22 (0.79 ; 1.40)
VT carriage prevalence		
Adults vs. <5y	-0.06 (-0.17; 0.05)	0.34 (0.05; 0.62)
5-17y vs. <5y	-0.10 (-0.41; 0.20)	0.73 (0.00; 1.43)
NVT carriage prevalence		
Adults vs. <5y	-0.10 (-0.30; 0.09)	0.76 (0.11; 1.46)
5-17y vs. <5y	-0.04 (-0.21; 0.14)	1.27 (0.63; 1.89)
VT proportion among carriers		
Adults vs. <5y	-0.07 (-0.41; 0.33)	0.73 (0.06;1.32)
5-17y vs. <5y	-0.03 (-0.41; 0.33)	0.73 (0.10; 1.37)

2.3.2. Age-dependent prevalence of VT carriage and distribution of VT serotypes

A total of eleven studies provided estimates of the distribution of VT and NVT serotypes in young children and in older children and/or adults. Estimates from a study in Israel [16] were based on outpatients, including a proportion of patients with upper respiratory tract infections (URTI), as serotype specific estimates could not be obtained for the healthy study participants only. Data on the serotype distribution among participants of two studies [18,19] were extracted from subsequent manuscripts [20,21] as the information was unavailable in the

original articles. Two studies included in the analysis only provided estimates in <2y olds rather than <5y olds. However, as meta-regression coefficients, obtained with and without the inclusion of such studies, were similar we included both studies in the final analysis.

Table 3 provides details of the studies included in the analysis. The proportion of VT serotypes isolated from positive swabs was consistently lower in older age groups compared to children under five, and that of NVT serotypes consistently higher. This finding was consistent across studies.

Table 3: Studies included in the Bayesian meta-regression of VT and NVT distribution, with proportion of VT out of positive samples, by age group

Studies	Country/Area	PCV valency	Age	VT/all (%VT) <5y olds	VT/all (%VT) 5-17y olds	VT/all (%VT) Adults	NT** in denom.
Hill PC.,(2006)[19] data in [21]	Gambia	PCV7 + 6A	<5y	108/197 (55%)	226/617 (37%)	186/674 (28%)	Yes
Adetifa IM., et al. (2012[34]	Nigeria	PCV7	<5y	173/375 (46%)	19/63(30%)	39/90 (43%)	Yes
Adetifa IM., et al. (2012) [34]	Nigeria	PCV10	<5y	174/375 (46%)	19/63(30%)	41/90 (45%)	Yes
Adetifa IM., et al. (2012) [34]	Nigeria	PCV13	<5y	264/375 (70%)	27/63(43%)	52/90 (58%)	Yes
Darboe et al (2012) [39]*, data in [58]	Gambia	PCV13	<1y	97/143 (68%)		6/26 (23%)	No
Mueller et al. (2012) [49]	Burkina Faso	PCV13	<5y	45/80 (56%)	20/57 (35%)	9/28 (32%)	No
Turner et al. (2012) [56]*	Thailand[1]	PCV13	<2y	105/188 (56%)		16/57 (28%)	No
van Gils E. et al (2009) [57]*	Netherlands	PCV7	<2y	115/213 (54%)		27/67 (40%)	Yes
Reis JN. et al. (2008)[54]	Brazil	PCV7	<5y	12/33 (36%)	12/43 (28%)	7/19 (37%)	Yes
Hammitt et al (2006) [3]	USA (Alaska)	PCV7	<5y	209/377 (55%)		78/275 (28%)	Yes
Regev-Yochay (2012) [52]	Occup. Palest Terr.[2]	PCV7	<5y	65/189 (34%)		6/30 (20%)	Yes
Regev-Yochay (2012) [52]	Occup. Palest Terr.	PCV10	<5y	69/189 (37%)		6/30 (20%)	Yes
Regev-Yochay (2012) [52]	Occup. Palest Terr.	PCV13	<5y	93/189 (49%)		9/30 (30%)	Yes
Regev-Yochay e(2004) [16]*	Israel	PCV7	<5y	87/200 (44%)		8/29 (28%)	No
Hussain (2005) [18] data in [20]	UK	PCV7	<5y	57/87 (66%)	8/15 (53%)	9/18 (50%)	No
Hussain (2005) [18] data in [20]	UK	PCV13	<5y	72/87 (83%)	11/15 (73%)	10/18 (55%)	No

[1]Maela refugee camp at the Thailand-Myanmar border. [2]Gaza strip. *data marked with (*) are based on approximation. **Non-typeable serotypes included in denominator (yes/no). See the Supporting Information section for further details.

We found a positive linear relationship between the proportion of VT serotypes isolated from carriers in children under five and that in 5-17y olds or in adults (Figure 3). The intercept was centred around zero in both models, and the slope of the coefficient was 0.73 (Credible intervals (CrI) 0.10 – 1.37) in 5-17y old and 0.73 (95%CrI 0.06; 1.32) in adult carriers compared to

the proportion of VT serotypes in children <5y. Further details on the model coefficients are provided in Table 2.

Figure 3: VT proportion in carried serotypes among <5y and older ages: scatter plot and fitted model
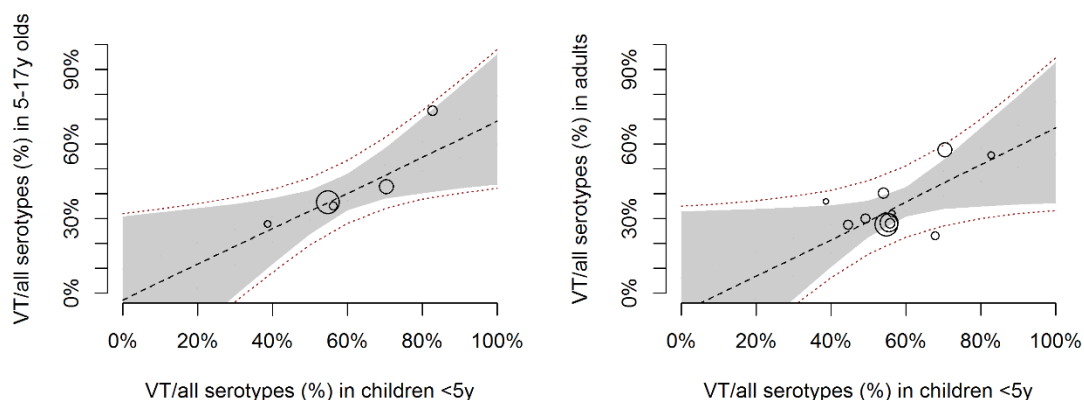
We then analysed the relationship between VT and NVT carriage prevalence in <5y olds and 5-17y olds or adults. As for the overall carriage prevalence, there was good evidence of a linear trend, with the prevalence of both VT and NVT in older age groups increasing with increasing VT and NVT prevalence in <5s (Figures 4 and 5).

However, given the shift in VT/NVT distribution in older age groups, the prevalence of VT serotypes in 5-17y olds and adults compared to that in children under five was comparatively lower than that of NVT serotypes (Table 3).

We performed subgroup analyses for VT serotypes included in PCV7 and VT serotypes included in PCV13, and found no difference in the coefficient obtained. Hence for studies reporting estimates for both PCV7 and PCV13 we included in the final analysis estimates with VT serotypes included in the higher valency vaccine.

Figure 4: VT and NVT prevalence in 5-17y olds against <5y olds: scatter plot and fitted model
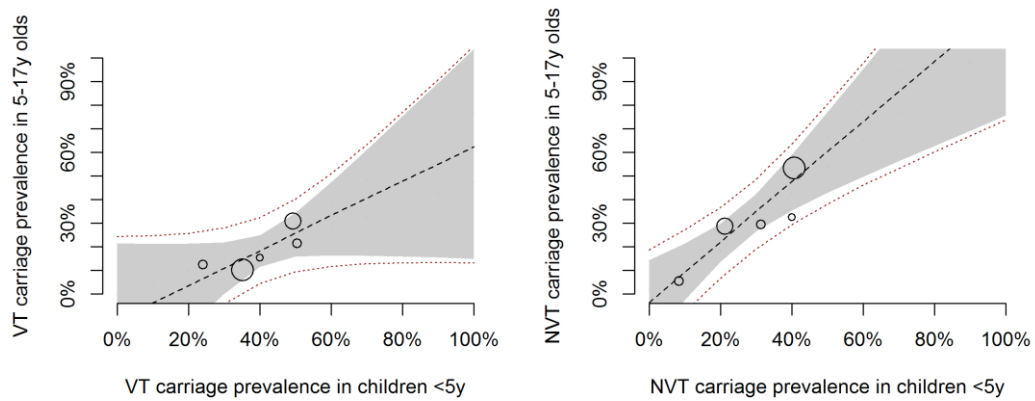
Legend for Figures 3 and 4: Each circle corresponds to one study, with the circle size proportional to the study size (i.e. number of individuals contributing to the x and y axis estimates). The lines correspond to the fitted Bayesian linear meta-regression model. The dashed black line shows the median posterior estimate and the grey shaded area the 95% credible interval around the median. The red dotted lines represent the 95% prediction interval.

Figure 5: VT and NVT prevalence in adults against <5y olds: scatter plot and fitted model
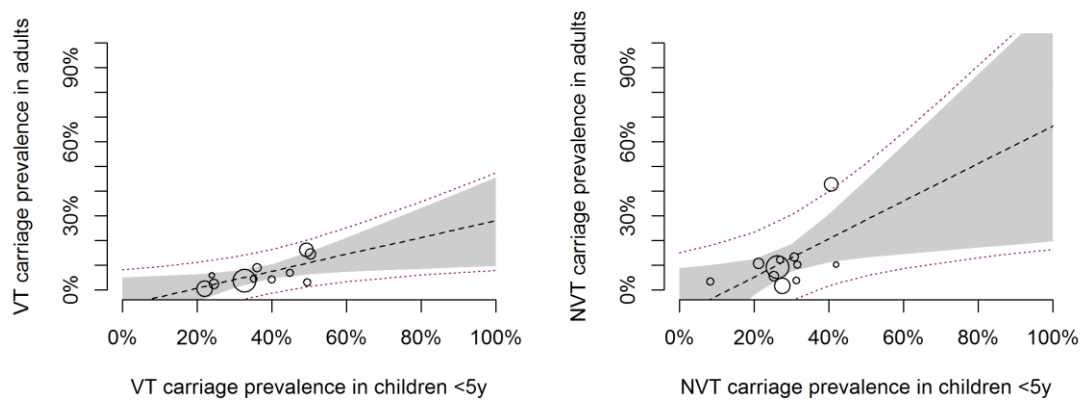
Each circle corresponds to one study, with the circle size proportional to the study size (i.e. number of individuals contributing to the x and y axis estimates). The lines correspond to the fitted Bayesian linear meta-regression model. The dashed black line shows the median posterior estimate and the grey shaded area the 95% credible interval around the median. The red dotted lines represent the 95% prediction interval

2.4. Discussion

In this systematic review and meta-analysis the prevalence of *Streptococcus pneumoniae* carriage in the nasopharynx of children aged under five was strongly correlated with the prevalence of nasopharyngeal carriage in older age groups. Furthermore, we found that the proportion of carriage attributed to vaccine serotypes was consistently decreasing with age. Our study provides a tool to help make informed predictions, however with some uncertainty, on the carriage prevalence and serotype distribution in older children and adults solely based on data in children <5y of age, which are more widely available.

The results of this study showed that despite the important geographic heterogeneity in carriage prevalence, there is a stable linear relationship between the carriage prevalence in young children and that in older children and adults. Such relationship also held for VT and NVT separately, although with different magnitude given the shift in serotype distribution in older age groups, with proportionally more NVT and less VT carriers. While a proportional decrease of carriage prevalence by age, as described by a linear correlation with an intercept centred around zero, did not describe the data well in most instances, the addition of a flexible intercept allowed for a good description of the age-dependent carriage association.

Although the decrease in carriage prevalence through childhood is a well-established fact, the between age group correlation in prevalence estimates has – to the best of our knowledge – not been previously described. These results are important to help improve our understanding of carriage and disease dynamics in the population, assess the potential population-wide effects of vaccination programmes and help design appropriate vaccination strategies. Given the high carriage prevalence rates found in children in many developing countries the indirect impact of routine infant PCV immunization on older children and adult populations in such

countries is likely to be high, as we find that carriage rates in those age groups are likely to high as well.

The general decrease in prevalence with increasing age can be caused by numerous factors, including the decrease in the duration of carriage with age [22,23], the reduction in the number of effective contacts as age increases, as well as the general maturation of the immune system [24].

Immunity induced by *S.pneumoniae* carriage is complex and still poorly understood [25]. Although carriage acquisition leads to the development of capsular antibodies, evidence suggests that such antibodies may not be the primary driver of the decrease in duration and prevalence of carriage with age [26]. Mouse models have shown that the development of immunity against colonization in mice depends on CD4+ T cells rather than serotype-specific antibodies as such, in particular T-cells secreting IL-17A   (T_H17 cells) [27], and there is evidence that T_H17 cells play a role in immunity against carriage in humans too [28].

Yet despite the uncertainty around the exact immune mechanisms, or the contribution of each of those towards acquired long term immunity, epidemiological evidence suggest that serotype-specific immunity against colonization is induced by acquisition of some serotypes such as 6A, 14 and 23F, which are included in PCV formulations and are some of the most prevalent serotypes in early childhood [24]. Hence the progressive acquisition of immunity against VT serotypes may also explain the shift in serotype distribution towards proportionally more NVT as age increases, as immunity against VT serotypes acquired in early childhood may reduce the likelihood of acquiring such serotypes later in life [29].

There are several direct applications of our study results.

Given that most studies are confined to children <5y of age only, the results of this study are particularly useful in the context of the progressive introduction of PCV10 or PCV13 in many

developing countries, in order to help estimate and appraise the possible impact of the vaccine across age groups. For example, nasopharyngeal carriage estimates are central to dynamic models of disease transmission [30,31], which can be used to model pre-vaccination dynamics and estimate post vaccination trends. With the results of this study, such models could be implemented in settings in which pre-vaccination data are only available in young children.

The quantification of the magnitude of change in carriage prevalence between children <5y and older age groups is also helpful in the sample size calculations of nasopharyngeal carriage surveys across age groups.

While the specific associations found between VT and NVT carriage prevalence across age groups may not hold for new higher valency vaccines under development, the estimates of overall carriage prevalence across age groups may help evaluate the possible population wide impacts on carriage of new protein-based or killed whole cell serotype-independent candidate vaccines [28].

As a practical application of the presented work we predicted the carriage prevalence in older children and adults based on carriage in under 5 year olds in an aboriginal population in the Northern Territory in Australia [33]. This study was not included in the analysis as it was conducted after the introduction of PCV. However, no change in overall carriage prevalence was observed in the three first years post PCV in this population. Hence we assumed that under a scenario of full serotype replacement overall carriage estimates by age group after PCV implementation would match those from the pre-PCV period. Using such data, we estimated the prevalence in 5-17y olds and adults to be 58.1% (95% prediction 38.7-77.4%) and 32.4% (95%prediction 14.8 – 48.8%) respectively, based on data in children under five. This closely matches the study estimates, with the prevalence in 5-17y olds estimated at 60.9% (95%CI 54.5% - 67.0%) and that in adults estimated at 26.0% (95%CI 22.3 – 29.9%).

Our study also suffers a range of limitations.

We did not restrict our analysis to any particular design and sampling strategy, and recruitment bias is likely to have occurred, in particular for studies based on convenience rather than random sampling. This is certainly the case for the adult age group, which in many studies consisted of the parents of children included in the study rather than a random cross sectional adult community sample. Such adults may therefore be more likely to be colonised, and to be colonised with a homologous strain to that in the children in the household, as shown in Chapter 4 of this thesis.

By restricting the analysis to broad age groups, we overlooked changes in carriage prevalence within each of those groups. In particular, the prevalence of carriage in the 5-17y olds is known to decline between the ages of 5 and 17 years. This may also account for some of the heterogeneity seen between studies, as in many studies the age representation of the 5-17y olds in the study sample may not have matched that of the general population.

Between-study heterogeneity may also have resulted from individual confounding factors associated with carriage prevalence, which we were unable to account for, such as malnutrition, antibiotic use or smoking [32,33].

Further, although standard WHO laboratory procedures [14] were reported in all studies, differences in swabbing techniques, number of colonies plated, processing of specimens and culture may also account for some of the differences seen.

We could not specifically estimate the carriage prevalence among the elderly as a function of that in young children, given the paucity of data. Having specific estimates in elderly would be important however, given the particularly high burden of pneumococcal disease in that age group and the potential indirect impact of routine PCV on carriage in them. While no significant difference in carriage prevalence between younger and older adults was reported in two studies that provided specific data on older adults or elderly [16,17], more data are required to enable specific estimates to be made for that age group.

Finally, model estimates of VT and NVT prevalence in older age groups as a function of that in young children were prone to more uncertainty than in the models based on overall carriage prevalence, given that fewer studies reported specific carriage data by groups of serotypes. In addition, those models were based on prevalence estimates which were mostly confined to the lower prevalence levels, resulting in substantial model uncertainty for high prevalence estimates. Further carriage studies will help improve the precision around such estimates, and the analysis can easily be updated with the model code provided in the Supporting Information.

2.5. Conclusions

Information on patterns of nasopharyngeal colonisation in individuals not directly targeted by pneumococcal conjugate vaccination is scarce but plays an important role in the consideration of the indirect impact of PCVs. We here present evidence that a non-trivial stable relationship between child and both adolescent and adult carriage rates holds. Furthermore, we show that a similar relationship for the proportion of vaccine type and non-vaccine type carriage is present. We exploit these and provide a tool to make an informed prediction of carriage rates in adolescents and adults based on childhood carriage rates only, including the associated uncertainty. Further carriage studies on broad age ranges will allow narrowing of the prediction intervals. If designed accordingly these could also provide the basis for childhood carriage informed estimates of carriage in the elderly population which is particularly affected by pneumococcal disease. An application of this model is undertaken in Research Paper 5 of the thesis.

Conflict of interest

None declared

References

1. Bogaert D, De Groot R, Hermans PW (2004) Streptococcus pneumoniae colonisation: the key to pneumococcal disease. Lancet Infect Dis 4: 144-154.

2. Halasa NB, Grijalva CG, Arbogast PG, Talbot TR, Craig AS, et al. (2013) Near Complete Elimination of the Seven Valent Pneumococcal Conjugate Vaccine Serotypes in Tennessee. Pediatr Infect Dis J.

3. Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, et al. (2006) Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. JInfectDis 193: 1487-1494.

4. Weinberger DM, Malley R, Lipsitch M (2011) Serotype replacement in disease after pneumococcal vaccination. Lancet 378: 1962-1973.

5. Fitzwater SP, Chandran A, Santosham M, Johnson HL (2012) The worldwide impact of the seven-valent pneumococcal conjugate vaccine. Pediatr Infect Dis J 31: 501-508.

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

7. Song JY, Choi JY, Lee JS, Bae IG, Kim YK, et al. (2013) Clinical and economic burden of invasive pneumococcal disease in adults: a multicenter hospital-based study. BMC Infect Dis 13: 202.

8. Griffin MR, Zhu Y, Moore MR, Whitney CG, Grijalva CG (2013) U.S. hospitalizations for pneumonia after a decade of pneumococcal vaccination. N Engl J Med 369: 155-163.

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

10. Melegaro A, Edmunds WJ (2004) Cost-effectiveness analysis of pneumococcal conjugate vaccination in England and Wales. Vaccine 22: 4203-4214.

11. Rozenbaum MH, van Hoek AJ, Fleming D, Trotter CL, Miller E, et al. (2012) Vaccination of risk groups in England using the 13 valent pneumococcal conjugate vaccine: economic analysis. BMJ 345: e6879.

12. Raftery AE, Li N, Sevcikova H, Gerland P, Heilig GK (2012) Bayesian probabilistic population projections for all countries. Proc Natl Acad Sci U S A 109: 13915-13921.

13. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, et al. (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 283: 2008-2012.

14. O'Brien KL, Nohynek H, World Health Organization Pneumococcal Vaccine Trials Carriage Working G (2003) Report from a WHO Working Group: standard method for detecting upper respiratory carriage of Streptococcus pneumoniae. Pediatr Infect Dis J 22: e1-11.

15. (2009) Cochrane Handbook for Systematic Reviews of Interventions; Higgins J, Green S, editors. Chichester, England: The Cochrane Collaboration and John Wiley & sons Ltd.

16. Regev-Yochay G, Raz M, Dagan R, Porat N, Shainberg B, et al. (2004) Nasopharyngeal carriage of Streptococcus pneumoniae by adults and children in community and family settings. Clin Infect Dis 38: 632-639.

17. Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JA (2008) The descriptive epidemiology of Streptococcus pneumoniae and Haemophilus influenzae nasopharyngeal carriage in children and adults in Kilifi district, Kenya. Pediatr Infect Dis J 27: 59-64.

18. Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. EpidemiolInfect 133: 891-898.

19. Hill PC, Akisanya A, Sankareh K, Cheung YB, Saaka M, et al. (2006) Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian villagers. Clin Infect Dis 43: 673-679.

20. Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoSMed 8: e1001017.

21. Roca A, Hill PC, Townend J, Egere U, Antonio M, et al. (2011) Effects of community-wide vaccination with PCV-7 on pneumococcal nasopharyngeal carriage in the Gambia: a cluster-randomized trial. PLoSMed 8: e1001107.

22. Melegaro A, Choi Y, Pebody R, Gay N (2007) Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. Am J Epidemiol 166: 228-235.

23. Melegaro A, Gay NJ, Medley GF (2004) Estimating the transmission parameters of pneumococcal carriage in households. EpidemiolInfect 132: 433-441.

24. Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, et al. (2008) Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. J Infect Dis 197: 1511-1518.

25. Miyaji EN, Oliveira ML, Carvalho E, Ho PL (2012) Serotype-independent pneumococcal vaccines. Cell Mol Life Sci.

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

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

28. Moffitt KL, Malley R (2011) Next generation pneumococcal vaccines. Curr Opin Immunol 23: 407-413.

29. Cobey S, Lipsitch M (2012) Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. Science 335: 1376-1380.

30. Choi YH, Jit M, Gay N, Andrews N, Waight PA, et al. (2011) 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PLoSOne 6: e26190.

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

32. Greenberg D, Givon-Lavi N, Broides A, Blancovich I, Peled N, et al. (2006) The contribution of smoking and exposure to tobacco smoke to Streptococcus pneumoniae and Haemophilus influenzae carriage in children and their mothers. Clin Infect Dis 42: 897-903.

33. Schaumburg F, Alabi A, von Eiff C, Flamen A, Traore H, et al. (2013) Streptococcus pneumoniae colonization in remote African Pygmies. Trans R Soc Trop Med Hyg 107: 105-109.

34. Adetifa IM, Antonio M, Okoromah CA, Ebruke C, Inem V, et al. (2012) Pre-vaccination nasopharyngeal pneumococcal carriage in a Nigerian population: epidemiology and population biology. PLoS One 7: e30548.

35. Bello Gonzalez T, Rivera-Olivero IA, Pocaterra L, Spadola E, Araque M, et al. (2010) [Pneumococcal carriage in mothers and children of the Panare Amerindians from the State of Bolivar, Venezuela]. Rev Argent Microbiol 42: 30-34.

36. Cekmez F, Karademir F, Tanju IA, Meral C, Pirgon O, et al. (2009) Pneumococcal serotypes recovered from health children and their possible association with risk factor in istanbul, Turkey. Int J Biomed Sci 5: 101-104.

37. Chen CJ, Huang YC, Su LH, Lin TY (2007) Nasal carriage of Streptococcus pneumoniae in healthy children and adults in northern Taiwan. Diagn Microbiol Infect Dis 59: 265-269.

38. Dagan R, Gradstein S, Belmaker I, Porat N, Siton Y, et al. (2000) An outbreak of Streptococcus pneumoniae serotype 1 in a closed community in southern Israel. Clin Infect Dis 30: 319-321.

39. Darboe MK, Thurnham DI, Morgan G, Adegbola RA, Secka O, et al. (2007) Effectiveness of an early supplementation scheme of high-dose vitamin A versus standard WHO protocol in Gambian mothers and infants: a randomised controlled trial. Lancet 369: 2088-2096.

40. Dhakal R, Sujatha S, Parija SC, Bhat BV (2010) Asymptomatic colonization of upper respiratory tract by potential bacterial pathogens. Indian J Pediatr 77: 775-778.

41. Granat SM, Mia Z, Ollgren J, Herva E, Das M, et al. (2007) Longitudinal study on pneumococcal carriage during the first year of life in Bangladesh. PediatrInfectDisJ 26: 319-324.

42. Greenberg D, Broides A, Blancovich I, Peled N, Givon-Lavi N, et al. (2004) Relative importance of nasopharyngeal versus oropharyngeal sampling for isolation of Streptococcus pneumoniae and Haemophilus influenzae from healthy and sick individuals varies with age. J Clin Microbiol 42: 4604-4609.

43. Henriqus Normark B, Christensson B, Sandgren A, Noreen B, Sylvan S, et al. (2003) Clonal analysis of Streptococcus pneumoniae nonsusceptible to penicillin at day-care centers with index cases, in a region with low incidence of resistance: emergence of an invasive type 35B clone among carriers. Microb Drug Resist 9: 337-344.

44. Inostroza J, Trucco O, Prado V, Vinet AM, Retamal G, et al. (1998) Capsular serotype and antibiotic resistance of Streptococcus pneumoniae isolates in two Chilean cities. Clin Diagn Lab Immunol 5: 176-180.

45. Kaltoft MS, Skov Sorensen UB, Slotved HC, Konradsen HB (2008) An easy method for detection of nasopharyngeal carriage of multiple Streptococcus pneumoniae serotypes. J Microbiol Methods 75: 540-544.

46. Leino T, Hoti F, Syrjanen R, Tanskanen A, Auranen K (2008) Clustering of serotypes in a longitudinal study of Streptococcus pneumoniae carriage in three day care centres. BMC Infect Dis 8: 173.

47. Lloyd-Evans N, O'Dempsey TJ, Baldeh I, Secka O, Demba E, et al. (1996) Nasopharyngeal carriage of pneumococci in Gambian children and in their families. Pediatr Infect Dis J 15: 866-871.

48. Lo WT, Wang CC, Yu CM, Chu ML (2003) Rate of nasopharyngeal carriage, antimicrobial resistance and serotype of Streptococcus pneumoniae among children in northern Taiwan. J Microbiol Immunol Infect 36: 175-181.

49. Mueller JE, Yaro S, Ouedraogo MS, Levina N, Njanpop-Lafourcade BM, et al. (2012) Pneumococci in the African meningitis belt: meningitis incidence and carriage prevalence in children and adults. PLoS One 7: e52464.

50. Nunes MC, Shiri T, van Niekerk N, Cutland CL, Groome MJ, et al. (2013) Acquisition of Streptococcus pneumoniae in pneumococcal conjugate vaccine-naive South African children and their mothers. Pediatr Infect Dis J 32: e192-205.

51. Parry CM, Diep TS, Wain J, Hoa NT, Gainsborough M, et al. (2000) Nasal carriage in Vietnamese children of Streptococcus pneumoniae resistant to multiple antimicrobial agents. Antimicrob Agents Chemother 44: 484-488.

52. Regev-Yochay G, Abullaish I, Malley R, Shainberg B, Varon M, et al. (2012) Streptococcus pneumoniae carriage in the Gaza strip. PLoS One 7: e35061.

53. Reichler MR, Allphin AA, Breiman RF, Schreiber JR, Arnold JE, et al. (1992) The spread of multiply resistant Streptococcus pneumoniae at a day care center in Ohio. J Infect Dis 166: 1346-1353.

54. Reis JN, Palma T, Ribeiro GS, Pinheiro RM, Ribeiro CT, et al. (2008) Transmission of Streptococcus pneumoniae in an urban slum community. J Infect 57: 204-213.

55. Sener B, Arikan S, Ergin MA, Gunalp A (1998) Rate of carriage, serotype distribution and penicillin resistance of Streptococcus pneumoniae in healthy children. Zentralbl Bakteriol 288: 421-428.

56. Turner P, Turner C, Jankhot A, Helen N, Lee SJ, et al. (2012) A longitudinal study of Streptococcus pneumoniae carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. PLoSOne 7: e38271.

57. van Gils EJ, Veenhoven RH, Hak E, Rodenburg GD, Bogaert D, et al. (2009) Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. JAMA 302: 159-167.

58. Darboe MK, Fulford AJ, Secka O, Prentice AM (2010) The dynamics of nasopharyngeal streptococcus pneumoniae carriage among rural Gambian mother-infant pairs. BMC Infect Dis 10: 195

2.6.   Supporting information

2.6.1.   PRISMA checklist

The results are reported in accordance with the PRISMA checklist, the full details of which can

be easily access online in the published version of the manuscript.

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0086136

 As per checklist, a structured summary was provided:

1. The introduction included (i) a clear rationale in the context of what is already known, (ii) a specific statement about the question being addressed with reference to with reference to participants, interventions, comparisons, outcomes, and study design (PICOS);

2. The methods section included (i) a description of the study protocol, (ii) the eligibility criteria for inclusion, (iii) the data sources for information, (iv) the full electronic search strategy, (v) the process of selecting study, (vi) the methods for data extraction, (vii) the list of variables for which data were sought, (viii) the methods used to describe the risk of bias in individual studies, (ix) a statement of the primary summary measures, and (x) a description of the methods used to handle data and combine data from different studies, including measures of consistency, (xi) a description of any assessment conducted to reduce the risk of bias and (xii) any additional analyses such as sensitivity analyses;

3. The results section included (i) an overview of the number of studies screened, assessed for eligibility and included in the review, with reasons for exclusion at each stage, (ii) a description of which data were extracted, (iii) a presentation of the risk of bias for each study and, if available, any outcome level assessment, (iv) For all outcomes considered (benefits or harms), present, for each study a summary data and effect

estimates and confidence intervals, ideally with a forest plot, (v) a presentation of the results of each meta-analysis done, including confidence intervals and measures of consistency, (vi) a presentation of any assessment of risk of bias across studies and (vii) results of any additional analyses, if done;

4. We provided a discussion section which (i) summarised the main findings, including strength of evidence of the outcome, (ii) discussed the main limitations of the study and (iii) provided a general interpretation of the results in the context of other evidence, and implications for future research.

5. Finally, the PRISMA checklist requires to describe all sources of funding and other support, which was done.

2.6.2. Details about data extraction from the respective studies included

1. **Abdullahi et al. [1]** : Estimates were directly provided in the paper.

2. **Adetifa et al [2]:** Overall carriage estimates were obtained based on the reported prevalence by age group and the number of individuals included in each age group. Estimates for PCV7, PCV10 and PCV13 serotypes were provided directly in the paper

3. **Bello Gonzales et al. [3]:** Estimates were directly provided in the paper.

4. **Cekmez et al [4]:**Estimates were directly provided in the paper.

5. **Chen et al. [5]:**Estimates were directly provided in the paper.

6. **Dagan et al. [6]:**Estimates were directly provided in the paper.

7. **Darboe et al.[7]**: We used data published in a subsequent study [8] as the latter provided more detailed information on carriage than the original study. A birth cohort was followed up to the age of one year, and nasopharyngeal (NP) samples were taken at 0, 2, 5, and 12 months. We calculated the carriage prevalence as the average prevalence between 0-12 months, based on monthly estimates. The carriage prevalence for months for which no NP sample was taken was calculated as the average of the NP

prevalence of the two nearest months with available estimates. Similarly, NP samples from mothers were taken at their child's birth 2 months, 5 months and 12 months later. We considered the sample taken at birth to be more representative from the carriage prevalence in the general adult population than those taken during the first year, during which the prevalence in mothers increase as a result of close infant-mother interaction. Hence only the carriage prevalence obtained at birth were used for the analysis of the carriage prevalence in adults. For the calculation of VT and NVT carriage prevalence NT serotypes were excluded from the denominator

8. **Dhakal et al. [9]**: Carriage estimates were directly provided in the paper

9. **Granat et al. [10]:** Carriage estimates were directly provided in the paper

10. **Greenberg et al. [11]**: Carriage estimates were directly provided in the paper

11. **Hammitt et al.[12]**: Carriage estimates were directly provided in the paper

12. **Henriqus Normark et al.[13]**: Carriage estimates were directly provided in the paper

13. **Hill et al. [14]**: Carriage estimates were directly provided in the paper

14. **Hussain et al. [15]:** Prevalence estimates were based on the study data later reported by **Flasche et al [16]**, where the prevalence by age group was adjusted for multiple testing. Sample size by age group was estimated based on the prevalence estimates in [16] and the number of cases, by age group. For the calculation of VT and NVT carriage prevalence NT serotypes were excluded from the denominator

15. **Inostroza et al. [17]:** The samples were obtained from two cities in two different populations: (1) Santiago, where children with Invasive Pneumococcal Disease (IPD) were sampled as well as their contacts, in addition to a sample of healthy children attending day care centres, and (ii) Temuco, where healthy children attending day care centre and hospital staff were sampled. For analysis, only samples from Temuco were considered.

16. **Kaltoft et al.[18]:** Carriage estimates were directly provided in the paper

17. **Leino et al.[19]**: Carriage estimates were directly provided in the paper

18. **Lloyd-Evans et al.[20]**: The study included index cases <5 years of age admitted to the hospital with a diagnosis of IPD, as well as community based healthy controls <5 years of age. Samples of family members of index cases were obtained on the day of admission as well as one month later. We included data from samples obtained in healthy children <5y in the analysis, as well as data from the samples from family members of cases of IPD obtained one month after the index cases' hospital admission.

19. **Lo et al. [21]**: Estimates were directly provided in the paper.

20. **Mueller et al. [22]**: Overall carriage estimates obtained based on the reported prevalence by age group and the number of individuals included in each age group.

21. **Nunes et al [23]**: The study participants consisted of 120 pairs of HIV positive mothers and HIV exposed children and 123 HIV unexposed mother-infant pairs. Swabs were collected from infants at five different times (average age of 4m, 7m, 9m, 12m and 16m). The carriage prevalence in children and mothers was approximated by the number of positive swabs in each age group over the total number of swabs collected. This calculation was performed to obtain overall carriage estimates. Estimates for PCV7 and PCV13 serotypes could not be obtained. Only the data from the HIV unexposed mother-infant pairs were used in this analysis.

22. **Parry et al. [24]**: Estimates were directly provided in the paper.

23. **Regev-Yochay et al. [25]**: The study population comprised of children and adults attending any of the health care centres in four large cities in Israel for any reason. Of the 404 children included, 173 attended for an upper respiratory tract infection (URTI) and among adults, 192 of the 1300 included had an URTI on the day. For the purpose of the analysis of overall carriage prevalence, children and adults with an URTI were excluded. However in the analysis of VT and NVT prevalence and distribution all data

were included as data available from the manuscript did not differentiate between those admitted or not for URTI.

24. **Regev-Yochay et al. [26]**: Overall carriage estimates obtained based on the reported prevalence by age group and the number of individuals included in each age group.

25. **Reichler et al.[27]**: Exact numbers were obtained based on a bar chart of the prevalence by age group and the number of individuals in each age group. We included all children <64months in the calculation of the carriage prevalence in <5y olds and children in the after kindergarten/after school aged 5-10y in the older children age groups. Estimates from a group of children aged 3-7y who attended a Montessori were not included as the age band stretched over the 2 age groups.

26. **Reis et al. [28]**: Carriage estimates were directly provided in the paper

27. **Sener et al. [29]:** Carriage estimates were directly provided in the paper

28. **Turner et al. [30]**: A cohort of newborns and their mothers were followed from birth to 24 months of age, and monthly swabs were taken. After a rapid increase in NP carriage prevalence in the first 2 months the prevalence of carriage in children remained stable between the ages of 3 and 24 months. The carriage prevalence was also stable in mothers. We calculated the carriage prevalence as the total number of positive swabs out of all swabs taken (children: 3363/4191 and mothers 1033/4195), and translated this in number of cases based on the sample size of children (n=234) and mothers (n=231) and the prevalence estimates. Given the stability of the positivity rate by monthly age group, we assumed for the analysis that the prevalence in <1y was equivalent to that in <2y.Similarly, for VT and NVT serotypes, the calculation was based on the average VT and NVT positive swabs (excluding NT) and the numbers of carriers calculated in each age group.

29. **van Gils et al.[31]**: An individual RCT was conducted. We used carriage estimates from the control group only (i.e. unvaccinated). Children were swabbed at 12 months, 18

months and 24 months. For carriage estimates in infants (<1y) we only included estimates at 12 months of age. However, for the VT and NVT carriage prevalence we calculated the average over the three sampling periods, at 12 months, 18 months and 24 months. In adults swabs were taken from parents when their children were 12 months and 24 months old. We calculated the carriage prevalence based on the average of samples taken at 12 months and samples taken at 24 months.

References

1. Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JA (2008) The descriptive epidemiology of Streptococcus pneumoniae and Haemophilus influenzae nasopharyngeal carriage in children and adults in Kilifi district, Kenya. Pediatr Infect Dis J 27: 59-64.

2. Adetifa IM, Antonio M, Okoromah CA, Ebruke C, Inem V, et al. (2012) Pre-vaccination nasopharyngeal pneumococcal carriage in a Nigerian population: epidemiology and population biology. PLoS One 7: e30548.

3. Bello Gonzalez T, Rivera-Olivero IA, Pocaterra L, Spadola E, Araque M, et al. (2010) [Pneumococcal carriage in mothers and children of the Panare Amerindians from the State of Bolivar, Venezuela]. Rev Argent Microbiol 42: 30-34.

4. Cekmez F, Karademir F, Tanju IA, Meral C, Pirgon O, et al. (2009) Pneumococcal serotypes recovered from health children and their possible association with risk factor in istanbul, Turkey. Int J Biomed Sci 5: 101-104.

5. Chen CJ, Huang YC, Su LH, Lin TY (2007) Nasal carriage of Streptococcus pneumoniae in healthy children and adults in northern Taiwan. Diagn Microbiol Infect Dis 59: 265-269.

6. Dagan R, Gradstein S, Belmaker I, Porat N, Siton Y, et al. (2000) An outbreak of Streptococcus pneumoniae serotype 1 in a closed community in southern Israel. Clin Infect Dis 30: 319-321.

7. Darboe MK, Thurnham DI, Morgan G, Adegbola RA, Secka O, et al. (2007) Effectiveness of an early supplementation scheme of high-dose vitamin A versus standard WHO protocol in Gambian mothers and infants: a randomised controlled trial. Lancet 369: 2088-2096.

8. Darboe MK, Fulford AJ, Secka O, Prentice AM (2010) The dynamics of nasopharyngeal streptococcus pneumoniae carriage among rural Gambian mother-infant pairs. BMC Infect Dis 10: 195.

9. Dhakal R, Sujatha S, Parija SC, Bhat BV (2010) Asymptomatic colonization of upper respiratory tract by potential bacterial pathogens. Indian J Pediatr 77: 775-778.

10. Granat SM, Mia Z, Ollgren J, Herva E, Das M, et al. (2007) Longitudinal study on pneumococcal carriage during the first year of life in Bangladesh. PediatrInfectDisJ 26: 319-324.

11. Greenberg D, Broides A, Blancovich I, Peled N, Givon-Lavi N, et al. (2004) Relative importance of nasopharyngeal versus oropharyngeal sampling for isolation of Streptococcus pneumoniae and Haemophilus influenzae from healthy and sick individuals varies with age. J Clin Microbiol 42: 4604-4609.

12. Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, et al. (2006) Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. JInfectDis 193: 1487-1494.

13. Henriqus Normark B, Christensson B, Sandgren A, Noreen B, Sylvan S, et al. (2003) Clonal analysis of Streptococcus pneumoniae nonsusceptible to penicillin at day-care centers with index cases, in a region with low incidence of resistance: emergence of an invasive type 35B clone among carriers. Microb Drug Resist 9: 337-344.

14. Hill PC, Akisanya A, Sankareh K, Cheung YB, Saaka M, et al. (2006) Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian villagers. Clin Infect Dis 43: 673-679.

15. Hussain M, Melegaro A, Pebody RG, George R, Edmunds WJ, et al. (2005) A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. EpidemiolInfect 133: 891-898.

16. Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, et al. (2011) Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoSMed 8: e1001017.

17. Inostroza J, Trucco O, Prado V, Vinet AM, Retamal G, et al. (1998) Capsular serotype and antibiotic resistance of Streptococcus pneumoniae isolates in two Chilean cities. Clin Diagn Lab Immunol 5: 176-180.

18. Kaltoft MS, Skov Sorensen UB, Slotved HC, Konradsen HB (2008) An easy method for detection of nasopharyngeal carriage of multiple Streptococcus pneumoniae serotypes. J Microbiol Methods 75: 540-544.

19. Leino T, Hoti F, Syrjanen R, Tanskanen A, Auranen K (2008) Clustering of serotypes in a longitudinal study of Streptococcus pneumoniae carriage in three day care centres. BMC Infect Dis 8: 173.

20. Lloyd-Evans N, O'Dempsey TJ, Baldeh I, Secka O, Demba E, et al. (1996) Nasopharyngeal carriage of pneumococci in Gambian children and in their families. Pediatr Infect Dis J 15: 866-871.

21. Lo WT, Wang CC, Yu CM, Chu ML (2003) Rate of nasopharyngeal carriage, antimicrobial resistance and serotype of Streptococcus pneumoniae among children in northern Taiwan. J Microbiol Immunol Infect 36: 175-181.

22. Mueller JE, Yaro S, Ouedraogo MS, Levina N, Njanpop-Lafourcade BM, et al. (2012) Pneumococci in the African meningitis belt: meningitis incidence and carriage prevalence in children and adults. PLoS One 7: e52464.

23. Nunes MC, Shiri T, van Niekerk N, Cutland CL, Groome MJ, et al. (2013) Acquisition of Streptococcus pneumoniae in pneumococcal conjugate vaccine-naive South African children and their mothers. Pediatr Infect Dis J 32: e192-205.

24. Parry CM, Diep TS, Wain J, Hoa NT, Gainsborough M, et al. (2000) Nasal carriage in Vietnamese children of Streptococcus pneumoniae resistant to multiple antimicrobial agents. Antimicrob Agents Chemother 44: 484-488.

25. Regev-Yochay G, Raz M, Dagan R, Porat N, Shainberg B, et al. (2004) Nasopharyngeal carriage of Streptococcus pneumoniae by adults and children in community and family settings. Clin Infect Dis 38: 632-639.

26. Regev-Yochay G, Abullaish I, Malley R, Shainberg B, Varon M, et al. (2012) Streptococcus pneumoniae carriage in the Gaza strip. PLoS One 7: e35061.

27. Reichler MR, Allphin AA, Breiman RF, Schreiber JR, Arnold JE, et al. (1992) The spread of multiply resistant Streptococcus pneumoniae at a day care center in Ohio. J Infect Dis 166: 1346-1353.

28. Reis JN, Palma T, Ribeiro GS, Pinheiro RM, Ribeiro CT, et al. (2008) Transmission of Streptococcus pneumoniae in an urban slum community. J Infect 57: 204-213.

29. Sener B, Arikan S, Ergin MA, Gunalp A (1998) Rate of carriage, serotype distribution and penicillin resistance of Streptococcus pneumoniae in healthy children. Zentralbl Bakteriol 288: 421-428.

30. Turner P, Turner C, Jankhot A, Helen N, Lee SJ, et al. (2012) A longitudinal study of Streptococcus pneumoniae carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. PLoSOne 7: e38271.

31. van Gils EJ, Veenhoven RH, Hak E, Rodenburg GD, Bogaert D, et al. (2009) Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. JAMA 302: 159-167.

3. Research paper 2: The efficacy and duration of protection of pneumococcal conjugate vaccines against nasopharyngeal carriage: a meta-regression model

Olivier le Polain de Waroux[1]; Stefan Flasche[1]; David Prieto-Merino[2]; David Goldblatt[3], W. John Edmunds[1]

1. Centre for the mathematical modelling of infectious diseases, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, WCE1 7HT, London, United Kingdom

2. Clinical Trials Unit, London School of Hygiene and Tropical Medicine, WCE1 7HT, London, United Kingdom

3. Institute of Child Health, University College London, WC1N 1EH, London, United Kingdom

LONDON
SCHOOL *of*
HYGIENE
&TROPICAL
MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED <u>FOR EACH</u> RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Olivier le Polain de Waroux
Principal Supervisor	W John Edmunds
Thesis Title	Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	Pediatric Infectious Disease Journal (PIDJ)		
When was the work published?	2015		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	NA
Please list the paper's authors in the intended authorship order:	NA
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate conceived the project, designed the study, explored statistical methods to address the question. Input was provided by co-authors in the design and statistical approach. The candidate compiled the data, performed the analysis and wrote the first draft of the manuscript. All co-authors

	reviewed the manuscript and approved its final version

Student Signature: _____ **Date:** 18 November 2016

Supervisor Signature: W.JE_____ **Date:** 18 November 2016

Evidence from the publisher

>>> "Braithwait, Celia" <Celia.Braithwait@wolterskluwer.com> 16/11/2016 06:14 >>>

Dear Olivier,

Yes, you are correct. If you were to continue through with your order and select the QUICK PRICE button, you should receive the following message:

This reuse is free of charge. No permission letter is needed from Wolters Kluwer Health, Lippincott Williams & Wilkins. We require that all authors always include a full acknowledgement. Example: AIDS: 13 November 2013 - Volume 27 - Issue 17 - p 2679-2689. Wolters Kluwer Health Lippincott Williams & Wilkins© No modifications will be permitted.

Thank you for checking with the journal.

Best,
Celia

Abstract

Background

Pneumococcal conjugate vaccines (PCVs) reduce disease largely through their impact on nasopharyngeal (NP) carriage acquisition of *S.pneumoniae*, a precondition for developing any form of pneumococcal disease. We aimed to estimate the vaccine efficacy (VE_C) and duration of protection of PCVs against *S.pneumoniae* carriage acquisition, through meta-regression models.

Methods

We identified intervention studies providing nasopharyngeal carriage estimates among vaccinated and unvaccinated children at any time after completion of a full vaccination schedule. We calculated VE_C for PCV7 serotypes, grouped as well as individually, and explored cross-protective efficacy against 6A. Efficacy estimates over time were obtained using a Bayesian meta-logistic regression approach, with time since completion of vaccination as a covariate.

Results

We used data from 22 carriage surveys (15 independent studies) from 5 to 64 months after the last PCV dose, including 14,298 children. The aggregate VE_C for all PCV7 serotypes six months after completion of the vaccination schedule was 57% (95%CrI 50 – 65%), varying by serotype from 38% (19F) to 80% (4). Our model provides evidence of sustained protection of PCVs for several years, with an aggregate VE_C of 42% (95%CrI 19 – 54) at 5 years, although the waning differed between serotypes. We also found evidence of cross-protection against 6A, with a VE_C of 39% six months after a complete schedule, decreasing to zero within five years post vaccination.

Conclusion

Our results suggest that PCVs confer lasting protection against acquisition of pneumococcal carriage of the seven studied serotypes, for several years after vaccination, albeit with differences across serotypes.

3.1.   Introduction

Pneumococcal conjugate vaccines (PCVs) reduce disease largely through their impact on nasopharyngeal (NP) carriage acquisition of *Streptococcus pneumoniae* (the pneumococcus), a precondition for developing any form of pneumococcal disease (1). Additionally PCVs reduce the progression to disease of VT carriers (1). The effect of PCV on carriage also drives the herd immunity effect of the vaccine in routine immunization, through a reduction in the transmission of vaccine serotypes (VT) in the community (2). Recently, emphasis has been put on the importance of carriage as a proxy measure for PCV impact assessments, and for using carriage as an additional and essential biomarker in the licensure pathway of new pneumococcal vaccines (3, 4).

A recent systematic review of the direct impact of PCVs by dosing schedules (5) showed consistent reductions in carriage of the serotypes targeted by the vaccine, including a few years after vaccination, with evidence favouring 3-dose schedules over fewer doses. However, systematic estimates of the efficacy of PCVs against carriage and the duration of protection conferred are lacking. Such estimates will help improve predictions about the likely impact of introducing the vaccine in routine immunization under different epidemiological scenarios. Estimates of the rate of waning efficacy are important not only to quantify the level of individual protection over time, but also the degree to which vaccinated children contribute to reducing community transmission as they age. Efficacy against carriage estimates also provide a benchmark against which new vaccines and vaccines under development can be evaluated (3). We studied the vaccine efficacy and duration of protection of pneumococcal vaccines against carriage, through meta-regression models.

3.2. Methods

3.2.1. Search strategy

We identified intervention studies reported by Fleming-Dutra et al. (5) in a recent systematic review of PCV vaccination schedules, which was based on data published between 1994 and September 2010, with post-hoc supplementation of studies published from 2011. We searched for any additional study published between 2011 and 31 May 2014 using a similar strategy as in (5), using EMBASE and MEDLINE databases. Details are provided in the Supporting Information.

3.2.2. Inclusion criteria

We considered the following initial criteria for inclusion: (i) intervention studies (ii) providing nasopharyngeal carriage estimates in vaccinated and unvaccinated children, (iii) with children vaccinated as per routine schedule, including three primary doses ('3+0' schedule) or at least two primary doses with a booster dose ('2+1' and '3+1' schedules). We further restricted our analysis to studies of either 7-valent, 10-valent or 13-valent licensed vaccines (PCV7, PCV10 and PCV13) or unlicensed vaccines (e.g. PCV9 and PCV11) linked to similar carrier proteins as licensed vaccines, including the Corynebacterium diphtheria toxin mutant 197 (CRM197), meningococcal outer membrane protein complex (OMPC) or the non-typeable Haemophilus Influenzae derived protein D (NTHi). Studies based on vaccines conjugated to other proteins or for which immunological equivalence is unclear (such as tetra- and penta-valent vaccines (6-8)) were not included.

Given that PCVs are not known to affect carriage clearance (9, 10), that the average duration of vaccine serotype (VT) carriage in infants and young children is somewhere around two months, but may vary by setting and serotype (11-14), and that 2-4 weeks are required for the antibody response to peak after vaccination, we excluded any data collected earlier than four months after complete vaccination, when the prevalence and serotype distribution was considered non-stationary, as detailed elsewhere (15, 16).

3.2.3. Data extraction

All but four studies were PCV7 trials, with three other trials based on PCV9 and one on PCV10. We extracted data on the group of PCV7 serotypes, as well as each individual PCV7 serotype (4, 6B, 9V, 14, 18C, 19F, 23F). We also extracted data on serotype 6A, one of the most common serotypes, which shares immunological traits with 6B but is not included in PCV7, PCV9 or PCV10, to explore possible cross-reactive protective efficacy. Other potential cross-reactive serotypes, such as 19A, were not studied, due to limited data.

3.2.4. Analysis

We defined the vaccine efficacy against carriage acquisition (VE_C) as the relative reduction in the rate of carriage acquisition among vaccinated compared to unvaccinated children, in trial conditions. Although acquisition events cannot directly be observed, it is possible to obtain a robust estimate of VE_C from cross sectional data based on 1 - OR (odds ratio), under general assumptions, with the OR defined as the odds of vaccination among the (group of) VT serotype(s) (henceforth, the 'target' group) to the odds of vaccination among those not carrying any VT (henceforth the 'reference' group) (15-17). Hence, in calculating the VE_C for each individual PCV7 serotype, we included in the target group all vaccinated and unvaccinated carriers of the particular serotype and in the reference group all non-vaccine serotype (NVT) carriers and non-carriers. Other VT were excluded from the serotype-specific analysis to account for vaccine-induced within-host changes in the pneumococcal flora, as explained elsewhere (15). We also excluded all VT serotypes from the analysis of VE_C against 6A. Similarly, in trials based on vaccines with higher valency than PCV7, data on the additional VT serotypes were excluded. Further details about the methods and assumptions underpinning the estimation of VE_C from cross-sectional data are described elsewhere (15-17). The analysis was based on summary data by (group of) serotype(s), rather than individual-level data.

We explored whether the proportion of carried VT serotypes out of all VT serotypes differed between studies, based on data in unvaccinated children, and used I^2 values to quantify heterogeneity (18).

We used a Bayesian logistic meta-regression model to estimate the aggregate and serotype-specific VE_C and its waning. In the model, for each study i,

$$\log(\frac{P_{Vi}^{R}}{1 - P_{Vi}^{R}}) = \alpha_i$$

$$\log(\frac{P_{Vi}^{T}}{1 - P_{Vi}^{T}}) = \alpha_i + \theta_i + \beta_{1*}\log(t_i)$$

, where P_{Vi}^{R} and P_{Vi}^{T} are the proportion of vaccinated individuals in the reference and target groups respectively, θ_i is the study-specific natural logarithm of the OR and β_1 represents the coefficient by which the log(OR) changes for each increase in the natural logarithm of time t since the peak VE_C (i.e. 4 months after vaccination), such that $\log(OR_i) = \theta_i + \beta_{1*}\log(t_i)$, with time in months.

We used a random effect model taking the between-study heterogeneity into account by assuming that θ_i were independent and sampled from a normal distribution centred around the mean log(OR) of carriage (μ) with a precision τ, such that $\theta_i \sim N(\mu, \tau)$ and $\tau = 1/\sigma^2$, where σ^2 is the between-study variance. A fixed effect was assumed for β_1.

The VE_C at time t can therefore be expressed as follows:

$$VEc_t = 1 - (e^\mu * t^{\beta_1})$$

We assigned uniform priors to α (unif (-10; 10)), μ (unif (-10, 0)), σ (unif (0,10)) and β_1 (unif (0,10)). The time coefficient β_1 was constrained to positive values, with the assumption that the efficacy should be declining. This assumption was further tested in a sensitivity analysis, by placing an unconstrained prior on β_1 (unif (-10,10)) .

Some studies provided more than one estimate. However, we did not adjust for the lack of independence due to the limited number of estimates from each study.

We explored the impact of schedule (booster (3+1 or 2+1) vs. non-booster (3+0)) by including schedule as a covariate in a multivariable model, and assigned a normal uninformed prior to its coefficient ($\beta_2 \sim N(0,10^3)$). We used an interaction term between schedule and time to look for a difference in the waning by schedule, with a normal uninformed prior on the interaction coefficient ($\beta_3 \sim N(0,10^3)$). Studies in which a 23-valent polysaccharide vaccine (PPV23) booster dose was provided after a primary schedule (as in (19, 20)) were considered part of the 3+0 group, given the lack of effect of PPV23 on carriage (19).

Finally, we conducted sensitivity analyses to explore the impact on our pooled VE_C estimates of omitting any one study. We also analysed two additional models of waning VE_C, including a model where time was included as a linear covariate and another model with an asymptotic function in which the VE_C of carriage approaches zero as time approaches infinity. Models were compared using the Deviance Information Criterion (DIC), a likelihood-based model fitting

statistic for Bayesian models similar to the frequentist Akaike Information Criterion (21). Further details are presented in the Supporting information.

Posterior distributions were obtained through a Markov Chain Monte Carlo (MCMC) Gibbs sampling algorithm based on 2 chains of 100,000 iterations running in parallel, after a burn-in of 5,000 iterations. The model was implemented in R using the jags package (22).

3.3.   Results

3.3.1.   Characteristics of the studies included

Of the eighteen intervention studies identified in (5), four were based on non-equivalent vaccines (6-8, 23) and one provided carriage data three months after the last dose (24), hence we ended up with thirteen studies. We identified two additional studies through our literature review, including a PCV7 trial with data six months after a 3+1 schedule (25), and another PCV10 trial with data collected in the first (26) as well the second year (27) after vaccination. Figure S1 (Supporting Information) shows the results of the literature search. Our analysis therefore included 15 individual publications (7, 10, 19, 20, 26-36) providing estimates from 22 different surveys, spanning from 5 months to 64 months after vaccination, and including 7,485 samples from vaccinated children, and 6,813 from unvaccinated children. All but four studies were based on PCV7. Three were PCV9 trials (29, 32, 37) and one was a PCV10 trial (26, 27). We were unable to restrict the latter to PCV7 serotypes only (as all data for PCV10 serotype were aggregated), and we explored the sensitivity of our model output to including (or not) data from that study. Nine data points were from surveys after booster vaccination (Table 4). Two studies (10, 33) were nested within a cluster randomized trial. The clustering was not adjusted for, and we explored the impact of those study estimates in the sensitivity analysis (see below). Serotype-specific data were obtained for 10 studies (7 PCV7 and 3 PCV9 studies), with 14 data points (10, 20, 25, 28, 29, 31-35).

Table 4: Studies included in the analysis

| Study | Country | PCV | Schedule | Time since last PCV dose | Number of children | | PCV7 VE$_c$ (95CrI) at each sample collection in the survey |
					Intervention	Control	
Cheung et al. (2009) [29]	The Gambia	PCV9	3+0	6 months, 16 months	1051,922	1041,925	58% (49%; 65%), 54% (44%; 62%)
Dagan et al. (2012) [28]	Israel	PCV7	3+0	6 months	329	170	51% (24%; 68%)
Kilpi et al. (2001)	Finland	PCV7	3+1	6 months	1603	800	41% (23%; 54%)
Lakshman et al. (2003) [20]	UK	PCV7	3+1PPV23	29 months, 36 months	150,143	126,188	29% (-49%; 66%), 5% (-59%; 43%)
Madhi et al. (2007) [31]	South Africa	PCV9	3+0	64 months	121	150	36% (-25%; 68%)
Mbelle et al. (1999) [32]	South Africa	PCV9	3+0	6 months	242	239	62% (43%; 76%)
Millar et al. (2006)	USA	PCV7	3+1	27 months	468	281	45% (14% ; 64%)
Obaro et al. (2000)	The Gambia	PCV7	3+0	5 months	98	99	45% (-2%; 70%)
O'Brien et al. (2007) [10]	USA	PCV7	3+0, 3+1	7.5 months (3+0), 7.5 months	226,239	208,219	65% (40%; 79%), 44% (11%; 65%)
Palmu et al. (2002) [34]	Finland	PCV7	3+1	46 months	401	353	46% (-16%; 76%)
Prymula et al. (2011) [26]	Czech Republic	PCV10*$	3+0, 3+1	8.5 months (3+0) ,	204	334	59% (27%; 79%)
				7 months (3+1), 12 months (3+1)	205,207	330,334	39% (-5%; 65%), 52% (12%; 75%)
Prymula et al. (2013) [27]	Czech Republic	PCV10*$	3+1	19 months	106	210	62% (19%; 83%)
Russell et al. (2010) [19]	Fiji	PCV7	3+1PPV23	6 months, 9 months	122,114	126,125	83% (53%; 95%), 70% (31%; 79%)
van Gils et al. (2009) [35]	The Netherlands	PCV7	2+1	7 months, 13 months	329,333	317,321	70% (56%; 79%), 70% (57%; 80%)
Yeh et al. (2003) [36]	USA	PCV7	3+0	6 months	41	28	-4% (-373%; 74%)

*In this trial two PCV10 arms were included, one receiving pre-vaccination paracetamol prophylaxis and one without prophylaxis. . Only data from the latter and the placebo group were included. $Serotype –specific data were not available and VEC in this trial is against all PCV10 serotypes, not PCV7 serotypes as in other studies included

3.3.2. Vaccine efficacy against carriage

We estimated a peak VE_C (i.e. 4 months after complete vaccination) of 62% (95%CrI 52 – 72%) against all VT serotypes, decreasing to 57% (95%CrI 50 – 65%) six months after vaccination, when the number of data points in the model is the highest, and 42% (95%CrI 19 – 54%) five years after vaccination (Figure 6, Table 5).

Figure 6: Plot of the model of VE_C over time and its 50% and 95% credible intervals, together with the individual study estimates
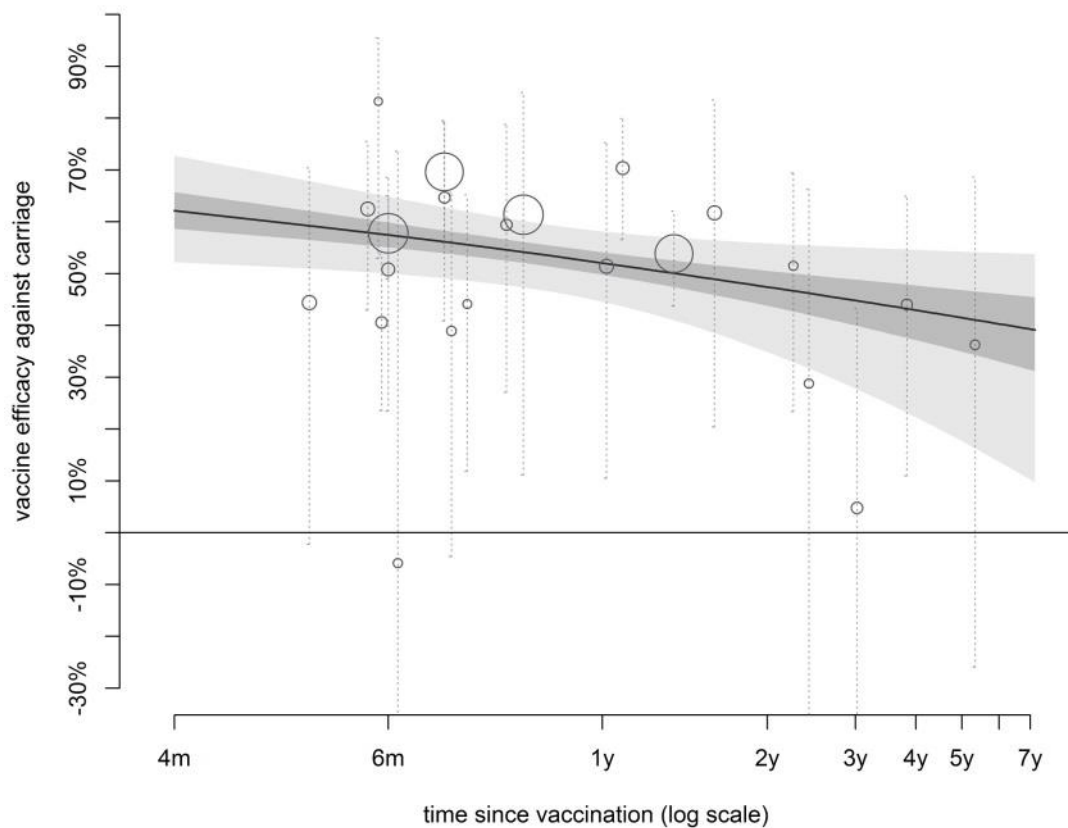
Legend: The plain line shows the model median, the dark grey shaded area the 50% credible interval (CrI) and the light grey shaded area the 95% CrI. The circles represent the point estimates of each individual study, with the size of the circle proportional to the study size, and the dotted vertical lines show the 95% confidence interval for each study. The horizontal axis represents time since complete vaccination

There was no evidence of a confounding effect of schedule on VE_C (with the coefficient β_2 centred around zero (-0.03 (95%CrI -0.32; 0.63)) or that the waning rate differed by schedule (interaction term β_3 0.01 (95%CrI -0.24; 0.13)). However, taken individually the median waning coefficient β_1 was smaller (i.e. 'flatter' slope) after a booster than after a 3+0 schedule (Figure 7 and Table 5). Our model was insensitive to the assumption on the prior of β_1.

Figure 7: The vaccine efficacy and its waning, for schedules with a booster (right panel) and without a booster dose (left panel).
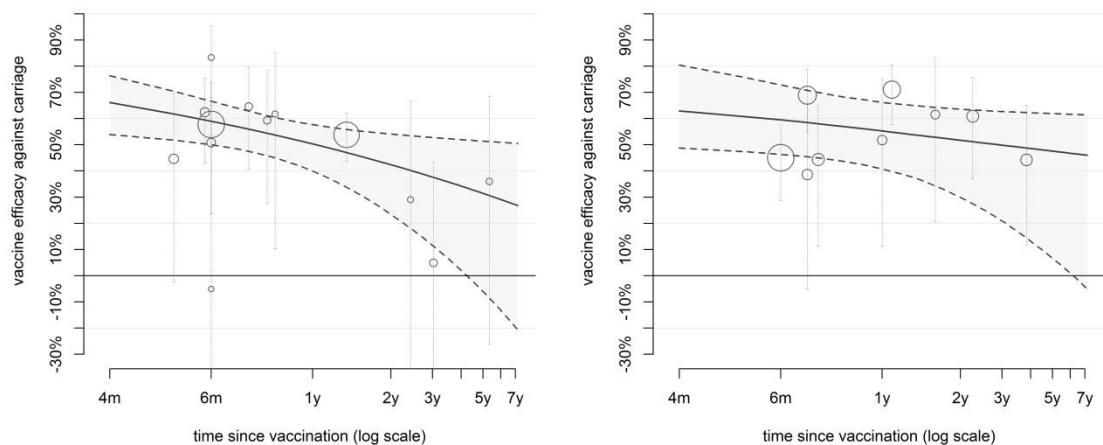
Legend: Left panel: model for 3+0 schedules. Left panel: 2+1 or 3+1 schedules. The plain dark regression line shows the model median, the grey shaded area the 95% credible interval (CrI) and the dotted lines the upper and lower bounds of the 95%CrI. The circles represent the point estimates of each individual study, with the size of the circle proportional to the study size, and the dotted vertical lines show the 95% confidence interval.

The serotype distribution among the unvaccinated children was fairly stable across studies (Figure 8), with little or moderate statistical heterogeneity in the distribution of serotypes among PCV7 positive samples (serotype-specific I^2 values of heterogeneity ranging from 0% to 60%). Serotypes 6B, 23F and 19F were the VT serotypes most commonly found, contributing to 26%, 22%, 28% respectively of the isolated PCV7 serotypes overall among unvaccinated children. Serotype 14 was found in 11% and serotypes 4, 9V and 18C in 3%, 6% and 3% of PCV7

samples. Serotype 6A was found in about 9% of unvaccinated children, a little higher than the prevalence of 6B (8%, p=0.07).

Figure 8: Distribution of the serotypes contained in PCV7, in each of the studies included in the serotype-specific model of vaccine efficacy
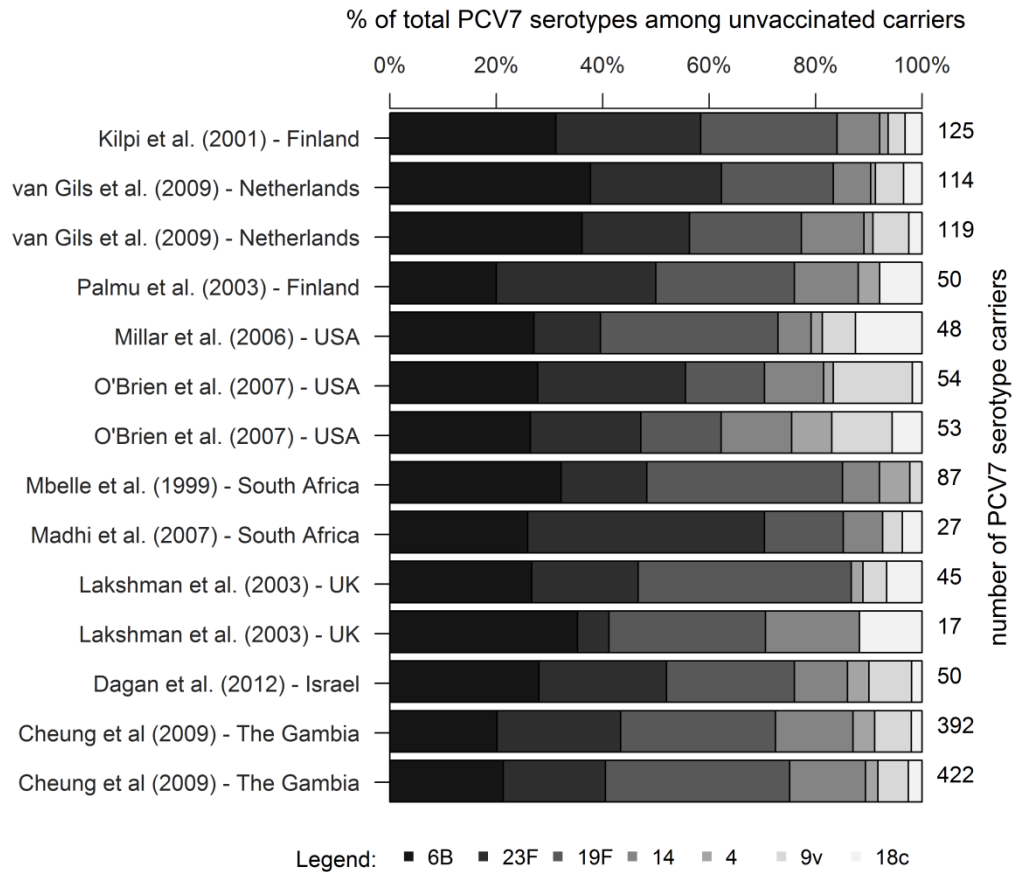
Efficacy estimates differed across PCV7 serotypes. Six months after vaccination the highest VE_C was measured for serotypes 4 (80%) and 9V (79%), and the lowest for 19F (38%) (Figure 9 and Table 5).

The decline in the efficacy over time varied by serotype (Table 5), with the slowest decline for serotypes 23F and 19F (median β_1 0.09) and more rapid declines for rarer serotypes, although credible intervals overlapped for all serotypes (Table 4).

We found evidence of protection against 6A, with a peak VE_C of 48% (95%CrI 18% – 72%), decreasing to zero within five years post vaccination (Table 5, Figure 9).

Figure 9: Serotype-specific models of vaccine efficacy against carriage, for each of the PCV7 serotypes as well as for serotype 6A
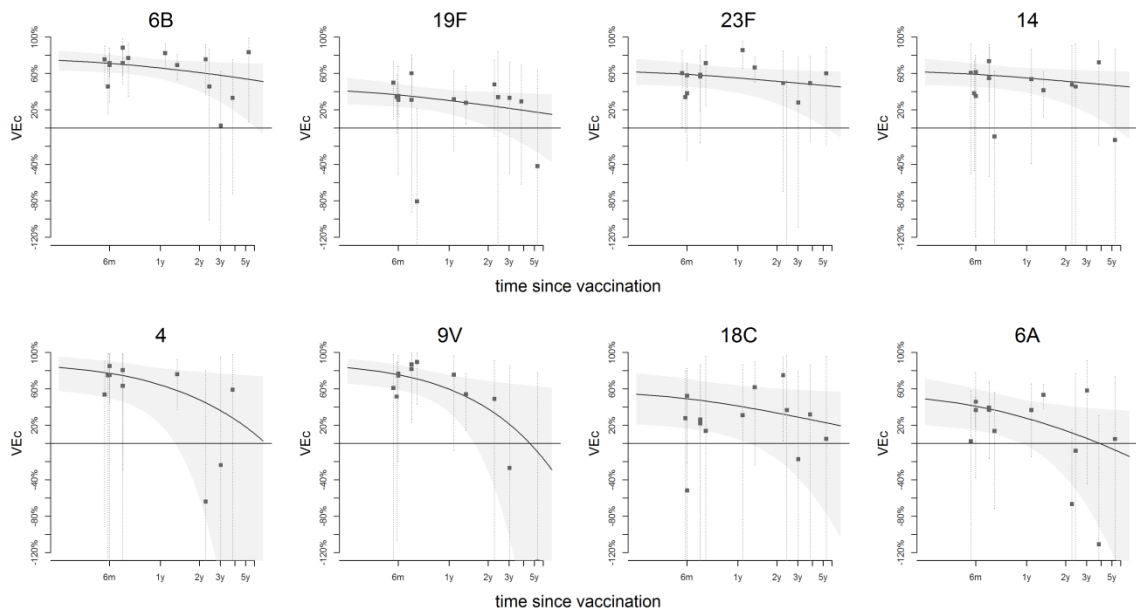
Legend: The black plain lines represent the model median and the grey shaded areas the model 95% credible interval. The squares and vertical dotted bars represent the study-specific point VEC estimates and their 95% confidence interval

Table 5: Aggregate and serotype-specific vaccine efficacy at different time points post vaccination, and model coefficient estimates

	VE$_C$ (95%CrI) at several time points after vaccination				Coefficient estimates (95%CrI)	
	Peak (4 months)	6 months	2 years	5 years	μ*	β$_1$ (waning)
PCV7 serotypes						
4	88% (62%;98%)	80% (54%;92%)	50% (-50%;78%)	18% (-328%;74%)	-2.11 (-3.75; -0.97)	0.46 (0.04; 1.21)
6B	77% (64%;89%)	72% (62%;83%)	62% (41%;72%)	54% (13%;71%)	-1.48 (-2.20; -1.03)	0.17 (0.01; 0.48)
9V	89% (71%;97%)	79% 64%;90%)	39% (-31%;69%)	-9% (-295%;62%)	-2.17 (-3.43; -1.24)	0.56 (0.11; 1.21)
14	64% (44%;81%)	57% (40%;71%)	40% (6%;58%)	29% (-44%;56%)	-1.01 (-1.65; -0.57)	0.16 (0.09; 0.47)
18C	59% (25%;82%)	52% (19%;73%)	34% (-14%;60%)	22% (-75%;58%)	-0.90 (-1.74; -0.29)	0.15 (0.01; 0.50)
19F	44% (28%;62%)	38% (24%;51%)	25% (3%;39%)	17% (-25%;37%)	-0.58 (-0.96; -0.33)	0.09 (0.01; 0.27)
23F	64% (49%;81%)	60% (46%;73%)	51% (25%;64%)	47% (3%;62%)	-1.02 (-1.64; -0.67)	0.09 (0.00; 0.36)
Cross reactive serotype						
6A	48% (18%%;72%)	39% (11%;58%)	16% (-33%%;41%)	0% (-95%;38%)	-0.65 (-1.28; -0.19)	0.15 (0.01; 0.43)
All PCV7 serotypes						
All schedules	62% (52%;72%)	57% (50%;65%)	47% (35%;56%)	42% (19%;54%)	-0.97 (-1.30;-0.72)	0.11 (0.01; 0.25)
Booster schedule	63% (49%;80%)	60% (47%;73%)	52% (30%;63%)	47% (6%;62%)	-1.00 (-1.64; -0.68)	0.08 (0.00; 0.36)
Primary dose schedule	66% (54%;77%)	59% (50%;67%)	42% (23%;54%)	31% (-7%;51%)	-1.10 (-1.49; -0.78)	0.18 (0.01; 0.37)

* mean log OR of carriage

3.4.   Sensitivity analysis

Our sensitivity analysis showed no significant impact of any study estimate on the coefficients. Estimates were similar after excluding the cluster randomized trial (10, 33), with a VE_C of 62% (51 – 73%) at 4 months, decreasing to 40% (12 – 54%) at 5 years. Excluding the two estimates from Cheung et al. in the Gambia (29), which together accounted for about 28% of all children included in the analysis, did not affect model estimates (VE_C of 62% (50 – 74%) at 4 months and 39% (12 – 54%) at 5 years). Finally, overall and booster schedule VE_C estimates and respective model coefficients were similar with and without data from the PCV10 trial (26, 27).

We explored two other models of waning, in addition to the main model (Supporting Information). In all three models there was good evidence of protective efficacy in the first few years after vaccination. A similar DIC was obtained for all three models estimating the aggregate VE_C, as well as for serotype-specific models, except for serotypes 14 and 19F for which the model with the asymptotic time function was outperformed by the other two. Further information can be found in the Supporting Information.

3.5.   Discussion

We computed pooled aggregate and PCV7 serotype-specific vaccine efficacy against nasopharyngeal acquisition and its waning based on a meta-regression model of cross-sectional data. Our results suggest that PCVs confers reasonable protection against acquisition of pneumococcal carriage of the seven studies serotypes, for several years after vaccination, albeit with differences across serotypes.

Previous studies have explored PCV efficacy against carriage (16) and compared schedules (5), however, a pooled estimate was not previously calculated. We found that the distribution of VT serotypes was relatively stable across settings, making the pooling of aggregate estimates

possible despite differences in the VE_C against individual serotypes. An analysis of the aggregate VE_C based on the pooled individual serotype-specific estimates also showed results in line with that of the main model presented, albeit with wider confidence bounds given the uncertainty around serotype-specific estimates.

Three serotypes (6B, 19F and 23F) accounted for about 75% of all PCV7 serotypes, but the VE_C for each of those differed, with high efficacy against 6B and a weaker anti-19F efficacy. A possible reason for this divergence is the difference in the amount of antibody required for protection as well as differences in the vaccine-induced opsonophagocytic activity (i.e. the ingestion and killing of pathogens by phagocytes), despite similar antibody geometric mean concentrations (GMCs) following PCV7 vaccination (30, 38). Interestingly, a recent study in the UK on the vaccine effectiveness and immune correlates of protection against IPD (39) showed that much less antibody is required for 6B and 23F protection than for 19F protection. The polysaccharide capsule of 19F is more resistant to complement deposition than 6B and requires higher levels of antibodies for opsonophagocytosis (38). However, although trials (37, 40, 41) have shown persistence of serum antibodies several years after vaccination, the exact mechanism underlying the protection against acquisition of carriage remains unclear. Such mechanisms could perhaps involve mucosal immunological responses (42) in addition to pre-existing circulating serum IgG, with serological markers incompletely capturing the mucosal response.

While natural immunity to colonization in infancy is poor, conjugate vaccines stimulate B-cell responses and the generation of memory B-cells (43), which can be naturally boosted. If boosting does contribute to maintaining a protective efficacy against carriage then one might expect efficacy to wane faster for rarer serotypes and slower for the more prevalent ones. Our results support such a hypothesis to some extent, showing a slower VE_C decline for the most

prevalent serotypes. This would also mean that VE_C may wane more rapidly after routine implementation of the vaccine than in trial conditions.

We also found evidence of cross-protective efficacy against 6A acquisition based on data from PCV7 and PCV9 trials. Such evidence is supported by trials and observational studies showing an impact of PCV7 on 6A disease (44), as well as immunological evidence with the vaccine eliciting functional antibody (i.e. antibodies inducing opsonophagocytosis) against 6A (45).

Efficacy estimates against carriage and their waning have several implications for vaccination programmes. Despite a stable distribution of serotypes across studies in this analysis, it is likely that some geographical variation occurs and serotype-specific efficacies are therefore important in predicting the impact of PCV in various epidemiological settings. Our results show good evidence of a direct protection against carriage in the first five years of life, when the pneumococcal burden is particularly high, and vaccinated children therefore also contribute to reducing transmission for several years. This may be particularly important in settings with low vaccine uptake or interrupted delivery.

The population impact on carriage and disease of introducing PCVs under different epidemiological scenarios could be explored through dynamic transmission models of pneumococcal disease, and our estimates of VE_C and their waning are essential parameters for such models.

The direct impact on disease is not solely conditioned on the VE_C, but also on the efficacy of the vaccine against progressing to disease following carriage (1). This explains the higher efficacy of PCV against invasive disease, at around 80% (1, 3, 39) . In contrast, the efficacy on disease progression against mucosal forms of disease, such as acute otitis media (AOM), is small with

most of the disease impact predicted by VE_C only (1, 3). Interestingly, the efficacy against pneumococcal AOM diagnosed by myringotomy with middle ear fluid aspiration among Finnish children enrolled in a large PCV7 trial (30) was 62% (48 – 72%) in the year following the booster dose, and serotype-specific efficacies were lowest for serotype 19F, at about 37%, and high for 6B (79%) 4 (75%,) and 9V (82%). Those estimates are similar to our aggregate and serotype-specific VE_C estimates, adding to the evidence that VE_C is a close measure of the efficacy against AOM.

An important question is the applicability of our results to 10- and 13-valent vaccines, given that many countries have introduced – or are planning to do so – those vaccines into their routine vaccination programmes. Data on immunological correlates of protection from trials generally suggest comparable responses after PCV13 and PCV10 vaccination compared to PCV7 (45, 48, 49). However, a recent study comparing IgG concentration and functional antibodies in PCV7 and PCV13 vaccinated Navajo and White Mountains Apache children in the US (50) found higher functional antibody activity against 19F after PCV13 vaccination, compared to PCV7, possibly explained by the inclusion of 19A in PCV13 and the additional activity of anti-19A antibodies against 19F. Another trial in the UK showed lower IgG concentrations for serotypes 4 and 23F for PCV13 compared to PCV7 recipients after a booster dose at 12 months of age. Those differences could translate in differences in aggregate VE_C.

The estimation of the efficacy against carriage acquisition from cross sectional data relies on several assumptions, the most important being that of stationarity – i.e. that the relationship between carriage incidence and carriage prevalence is stable (15, 16). Vaccination will introduce some temporary disturbance in the carriage rates of different serotypes, with the average prevalence estimates stabilising after some time (51). Auranen et al. (17) suggest that stationary levels should not be considered before at least twice the duration of carriage since vaccination.

We included studies from four months after vaccination to account for this, which we considered to be a good trade-off between ensuring steady-state carriage levels and avoiding peak estimates to be affected by waning VE_C.

The assumption that PCV do not affect clearance is based on limited evidence (9, 10). Similarly, studies have suggested that the vaccine may also impact carriage density (10). In both scenarios (reduced duration and reduced density), VE_C could represent a combined efficacy estimate against acquisition and transmission under the assumption that a reduction in duration of carriage and/or carriage density is associated with both a reduction in the likelihood of detection and of transmission, as discussed elsewhere (15, 16).

Our study has a number of additional limitations.

First, our analysis was limited by the number of data points, with wider uncertainty as time since vaccination increases and the smaller study sizes for serotype specific analyses, with substantial uncertainty around model estimates for the least prevalent serotypes. The small number of data points in each schedule subgroup may have limited our ability to detect any difference between schedules. In addition, we were unable to provide concise predictions about the waning later than five years after completion of the vaccination schedule. This was due to the lack of data and the absence of strong statistical evidence of the superiority of the main model of waning presented compared to the two other models tested.

Second, studies were based on the identification of the dominant serotype in single colonies and multiple colonization was not taken into account. If the prevalence of multiple colonization is low and if there are no differences in the propensity of detecting one serotype over another, VE_C estimates based on single colonization would nonetheless adequately capture VE_C (15).

There are several other factors related to vaccine schedules and delivery that may impact on VE_C (and on the heterogeneity between studies) which we were unable to explore, including the timing and spacing of doses and the co-administration of PCV with different childhood vaccines (48). For example, a recent systematic review of the impact of PCV vaccination schedules on immunological responses (48) suggests that immune responses to serotype 14 may be influenced by co-administration of PCV with DTP vaccines, with significantly higher GMCs observed with acellular pertussis compared to the whole cell pertussis vaccine.

In addition, although the description of the swabbing and sample processing techniques used in the studies included – although sometimes limited – seem to conform to WHO guidelines (52), we cannot rule out that some of the between-study heterogeneity may be due to differences in such techniques.

Finally, further research to obtain more precise estimates of VE_C after non-complete schedules, particularly single catch-up doses, is warranted. This is particularly relevant in the context of PCV roll out in low-income settings, as some countries may opt for catch-up campaigns at the introduction of the vaccine.

In conclusion, through this study we provide consistent evidence for a lasting efficacy against carriage acquisition of PCV in children during the first few years after completion of vaccination, although with differences in efficacy and duration of protection between serotypes.

Acknowledgements

Conflicts of interest

References

1. Simell B, Auranen K, Kayhty H, et al. The fundamental link between pneumococcal carriage and disease. *Expert review of vaccines*. 2012;11:841-855.

2. Davis S, Deloria-Knoll M, Kassa HT, O'Brien KL. Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: Review of evidence on indirect effects. *Vaccine*. 2013.

3. Goldblatt D, Ramakrishnan M, O'Brien K. Using the impact of pneumococcal vaccines on nasopharyngeal carriage to aid licensing and vaccine implementation; a PneumoCarr meeting report March 27-28, 2012, Geneva. *Vaccine*. 2013;32:146-152.

4. Weinberger DM, Bruden DT, Grant LR, et al. Using pneumococcal carriage data to monitor postvaccination changes in invasive disease. *American journal of epidemiology*. 2013;178:1488-1495.

5. Fleming-Dutra KE, Conklin L, Loo JD, et al. Systematic Review of the Effect of Pneumococcal Conjugate Vaccine Dosing Schedules on Vaccine-type Nasopharyngeal Carriage. *The Pediatric infectious disease journal*. 2014;33 Suppl 2:S152-160.

6. Dagan R, Muallem M, Melamed R, Leroy O, Yagupsky P. Reduction of pneumococcal nasopharyngeal carriage in early infancy after immunization with tetravalent pneumococcal

vaccines conjugated to either tetanus toxoid or diphtheria toxoid. *The Pediatric infectious disease journal.* 1997;16:1060-1064.

7. Obaro SK, Adegbola RA, Chang I, et al. Safety and immunogenicity of a nonavalent pneumococcal vaccine conjugated to CRM197 administered simultaneously but in a separate syringe with diphtheria, tetanus and pertussis vaccines in Gambian infants. *The Pediatric infectious disease journal.* 2000;19:463-469.

8. Dagan RZ, O. et al. "Nasopharyngeal (NP) carriage of Streptococcus pneumoniae (Pnc) in toddlers vaccinated during infancy with an 11 valent pneumococcal vaccine conjugated to diphtheria and tetanus toxoids (PCV-DT)." Abstracts of the Interscience Conference on Antimicrobial Agents & Chemotherapy 40: 236

2000.

9. Dagan R, Givon-Lavi N, Zamir O, et al. Reduction of nasopharyngeal carriage of Streptococcus pneumoniae after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers. *JInfectDis.* 2002;185:927-936.

10. O'Brien KL, Millar EV, Zell ER, et al. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. *The Journal of infectious diseases.* 2007;196:1211-1220.

11. Abdullahi O, Karani A, Tigoi CC, et al. Rates of acquisition and clearance of pneumococcal serotypes in the nasopharynges of children in Kilifi District, Kenya. *The Journal of infectious diseases.* 2012;206:1020-1029.

12. Lipsitch M, Abdullahi O, D'Amour A, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology.* 2012;23:510-519.

13. Turner P, Turner C, Jankhot A, et al. A longitudinal study of Streptococcus pneumoniae carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. *PLoSOne.* 2012;7:e38271.

14. Andrews N, Waight PA, Borrow R, et al. Using the indirect cohort design to estimate the effectiveness of the seven valent pneumococcal conjugate vaccine in England and Wales. *PloS one.* 2011;6:e28435.

15. Auranen K, Rinta-Kokko H, Halloran ME. Estimating strain-specific and overall efficacy of polyvalent vaccines against recurrent pathogens from a cross-sectional study. *Biometrics*. 2013;69:235-244.

16. Rinta-Kokko H, Dagan R, Givon-Lavi N, Auranen K. Estimation of vaccine efficacy against acquisition of pneumococcal carriage. *Vaccine*. 2009;27:3831-3837.

17. Auranen K, Rinta-Kokko H, Goldblatt D, et al. Colonisation endpoints in Streptococcus pneumoniae vaccine trials. *Vaccine*. 2013;32:153-158.

18. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003;327:557-560.

19. Russell FM, Carapetis JR, Satzke C, et al. Pneumococcal nasopharyngeal carriage following reduced doses of a 7-valent pneumococcal conjugate vaccine and a 23-valent pneumococcal polysaccharide vaccine booster. *Clinical and vaccine immunology : CVI*. 2010;17:1970-1976.

20. Lakshman R, Murdoch C, Race G, Burkinshaw R, Shaw L, Finn A. Pneumococcal nasopharyngeal carriage in children following heptavalent pneumococcal conjugate vaccination in infancy. *Archives of disease in childhood*. 2003;88:211-214.

21. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linden A. Bayesian measures of model complexity and fit. . *Journal of the Royal Statistical Society: Series B*. 2002;64.

22. JAGS 3.3.0 Last updated 26 April 2013. Available at: http://mcmc-jags.sourceforge.net/.

23. Watson K, Carville K, Bowman J, et al. Upper respiratory tract bacterial carriage in Aboriginal and non-Aboriginal children in a semi-arid area of Western Australia. *The Pediatric infectious disease journal*. 2006;25:782-790.

24. Prymula R, Kriz P, Kaliskova E, Pascal T, Poolman J, Schuerman L. Effect of vaccination with pneumococcal capsular polysaccharides conjugated to Haemophilus influenzae-derived protein D on nasopharyngeal carriage of Streptococcus pneumoniae and H. influenzae in children under 2 years of age. *Vaccine*. 2009;28:71-78.

25. Kilpi T, Syrjänen R, Palmu A, Herva E, Eskola J, Mäkelä P. Parallel evaluation of 7-valent PncCRM on pneumococcal carriage and AOM. . *European Society for Paediatric Infectious Diseases (ESPID) Conference*2001.

26. Prymula R, Hanovcova I, Splino M, et al. Impact of the 10-valent pneumococcal non-typeable Haemophilus influenzae Protein D conjugate vaccine (PHiD-CV) on bacterial nasopharyngeal carriage. *Vaccine*. 2011;29:1959-1967.

27. Usuf E, Bottomley C, Adegbola RA, Hall A. Pneumococcal carriage in sub-Saharan Africa--a systematic review. *PloS one*. 2014;9:e85001.

28. Dagan R, Givon-Lavi N, Porat N, Greenberg D. The effect of an alternative reduced-dose infant schedule and a second year catch-up schedule with 7-valent pneumococcal conjugate vaccine on pneumococcal carriage: a randomized controlled trial. *Vaccine*. 2012;30:5132-5140.

29. Cheung YB, Zaman SM, Nsekpong ED, et al. Nasopharyngeal carriage of Streptococcus pneumoniae in Gambian children who participated in a 9-valent pneumococcal conjugate vaccine trial and in their younger siblings. *The Pediatric infectious disease journal*. 2009;28:990-995.

30. Kilpi T, Ahman H, Jokinen J, et al. Protective efficacy of a second pneumococcal conjugate vaccine against pneumococcal acute otitis media in infants and children: randomized, controlled trial of a 7-valent pneumococcal polysaccharide-meningococcal outer membrane protein complex conjugate vaccine in 1666 children. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2003;37:1155-1164.

31. Madhi SA, Adrian P, Kuwanda L, Cutland C, Albrich WC, Klugman KP. Long-term effect of pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae--and associated interactions with Staphylococcus aureus and Haemophilus influenzae colonization--in HIV-Infected and HIV-uninfected children. *The Journal of infectious diseases*. 2007;196:1662-1666.

32. Mbelle N, Huebner RE, Wasas AD, Kimura A, Chang I, Klugman KP. Immunogenicity and impact on nasopharyngeal carriage of a nonavalent pneumococcal conjugate vaccine. *The Journal of infectious diseases*. 1999;180:1171-1176.

33. Millar EV, O'Brien KL, Watt JP, et al. Effect of community-wide conjugate pneumococcal vaccine use in infancy on nasopharyngeal carriage through 3 years of age: a cross-sectional study in a high-risk population. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2006;43:8-15.

34. Palmu AA, Verho J, PH M. Long-term efficacy of the seven valent PncCRM vaccine on nasopharyngeal carriage. . *3rd International Symposium on Pneumococci and Pneumcooccal Disease (ISPPD-3)*. Anchorage, Alaska2002.

35. van Gils EJ, Veenhoven RH, Hak E, et al. Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. *JAMA : the journal of the American Medical Association*. 2009;302:159-167.

36. Yeh SH, Zangwill KM, Lee H, et al. Heptavalent pneumococcal vaccine conjugated to outer membrane protein of Neisseria meningitidis serogroup b and nasopharyngeal carriage of Streptococcus pneumoniae in infants. *Vaccine*. 2003;21:2627-2631.

37. Madhi SA, Adrian P, Kuwanda L, et al. Long-term immunogenicity and efficacy of a 9-valent conjugate pneumococcal vaccine in human immunodeficient virus infected and non-infected children in the absence of a booster dose of vaccine. *Vaccine*. 2007;25:2451-2457.

38. Ekstrom N, Vakevainen M, Verho J, Kilpi T, Kayhty H. Functional antibodies elicited by two heptavalent pneumococcal conjugate vaccines in the Finnish Otitis Media Vaccine Trial. *Infection and immunity*. 2007;75:1794-1800.

39. Andrews NJ, Waight PA, Burbidge P, et al. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *The Lancet infectious diseases*. 2014.

40. Ekstrom N, Ahman H, Palmu A, et al. Concentration and high avidity of pneumococcal antibodies persist at least 4 years after immunization with pneumococcal conjugate vaccine in infancy. *Clinical and vaccine immunology : CVI*. 2013;20:1034-1040.

41. Truck J, Snape MD, Tatangeli F, et al. Pneumococcal serotype-specific antibodies persist through early childhood after infant immunization: follow-up from a randomized controlled trial. *PloS one*. 2014;9:e91413.

42. Zhang Q, Bernatoniene J, Bagrade L, et al. Regulation of production of mucosal antibody to pneumococcal protein antigens by T-cell-derived gamma interferon and interleukin-10 in children. *Infection and immunity*. 2006;74:4735-4743.

43. Kelly DF, Moxon ER, Pollard AJ. Haemophilus influenzae type b conjugate vaccines. *Immunology*. 2004;113:163-174.

44. Whitney CG, Pilishvili T, Farley MM, et al. Effectiveness of seven-valent pneumococcal conjugate vaccine against invasive pneumococcal disease: a matched case-control study. *Lancet*. 2006;368:1495-1502.

45. Esposito S, Tansey S, Thompson A, et al. Safety and immunogenicity of a 13-valent pneumococcal conjugate vaccine compared to those of a 7-valent pneumococcal conjugate vaccine given as a three-dose series with routine vaccines in healthy infants and toddlers. *Clinical and vaccine immunology : CVI*. 2010;17:1017-1026.

46. Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. *PLoSOne*. 2012;7:e39927.

47. Flasche S, Edmunds J, Millar E, Goldblatt D, Robertson C, Choi JY. The impact of specific and non-specific immunity on the ecology of Streptoccus pneumoniae and the implications for vaccination. *Proc R Soc B*. 2013.

48. Park DE, Johnson TS, Nonyane BA, et al. The differential impact of coadministered vaccines, geographic region, vaccine product and other covariates on pneumococcal conjugate vaccine immunogenicity. *The Pediatric infectious disease journal*. 2014;33 Suppl 2:S130-139.

49. Snape MD, Klinger CL, Daniels ED, et al. Immunogenicity and reactogenicity of a 13-valent-pneumococcal conjugate vaccine administered at 2, 4, and 12 months of age: a double-blind randomized active-controlled trial. *The Pediatric infectious disease journal*. 2010;29:e80-90.

50. Grant LR, O'Brien SE, Burbidge P, et al. Comparative immunogenicity of 7 and 13-valent pneumococcal conjugate vaccines and the development of functional antibodies to cross-reactive serotypes. *PloS one*. 2013;8:e74906.

51. Auranen K, Rinta-Kokko H, Goldblatt D, et al. Design questions for Streptococcus pneumoniae vaccine trials with a colonisation endpoint. *Vaccine*. 2013;32:159-164.

52. O'Brien KL, Moulton LH, Reid R, et al. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *Lancet*. 2003;362:355-361.

3.6. Supporting Information

3.6.1. Literature search to complement the existing systematic review

3.6.1.1. Search strategy

We searched for any additional study published between 2011 and 31 May 2014 using MEDLINE and EMBASE databases, and the same search strategy as in (5), but restricted to nasopharyngeal carriage as outcome.

We used the following keywords [all fields] :

Search #1: pathogen

"Streptococcus pneumoniae" OR ("Diplococcus" AND "pneumoniae") OR ("micrococcus" AND "pneumoniae") OR "Pneumococcus" OR "pneumococcal" OR "s.pneumoniae" OR "pneumococci" OR "streptococcus" OR "streptococcal" OR "Pneumococc"

Search #2: outcome

("Nasopharyngeal" AND "carriage") OR ("Nasopharyngeal" AND "colonization") OR ("Nasopharyngeal" AND "colonisation")

Search #3: vaccine

"Vaccine" OR "vaccines" OR "vaccination" OR "vaccinated" OR "immunization" OR "immunisation" OR "immunized" OR "immunised" OR "PCV" OR "Prevenar" OR "PCV7" OR "PCV-7" OR "PNCRM7" OR "PNCRM-7" OR "PCV10" OR "PCV-10" OR "PCV9" OR "PCV-9" OR "PCV11" OR "PCV-11".

3.6.1.2. Results

Combining those three searches yielded 468 citations. After automatic and manual de-duplication, we ended up with 208 citations to screen.

Of those, 179 were excluded based on the title or the abstract. The full text of 29 references were screened. Of those, three were from trials meeting our inclusion criteria, including a PCV7 trial from Israel (28) and a PCV10 trial from the Czech Republic, with two different publications (26, 27) (Figure S1 below). Additionally, we also retrieved data from a large Finnish trial presented at a conference in 2001 (25), and used illustratively by Auranen et al.(15)

Figure S1: Flow diagram of the literature search
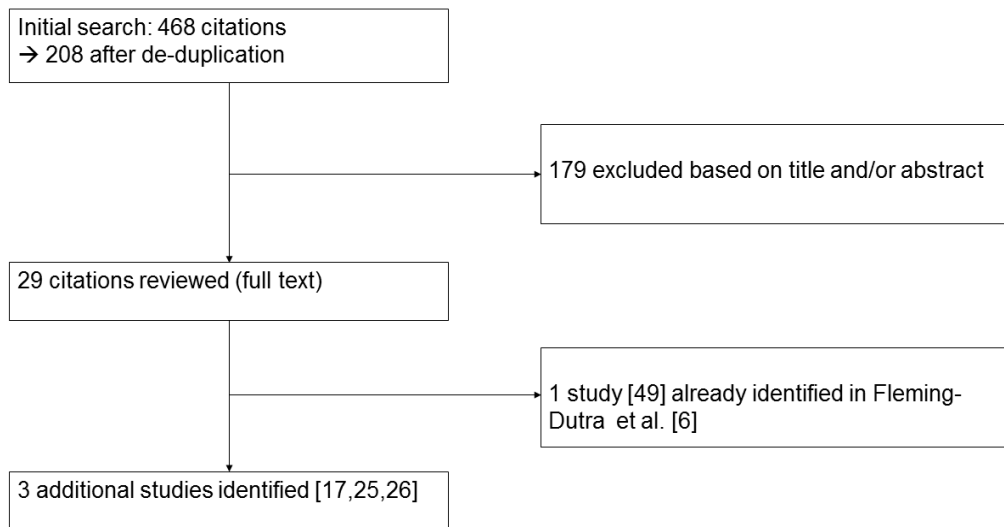
3.6.2. Comparing models of waning VE$_C$

Three models of waning VE$_C$ were considered.

For each study i,

$$\log(\frac{P_{Vi}^{R}}{1-P_{Vi}^{R}}) = \alpha_i \qquad \text{\# In all models}$$

$$\log(\frac{P_{Vi}^{T}}{1-P_{Vi}^{T}}) = \alpha_i + \theta_i + \beta_1 * \log(t_i) \qquad \text{\# In model 1 (main model presented)}$$

$$\log(\frac{P_{Vi}^{T}}{1-P_{Vi}^{T}}) = \alpha_i + \theta_i + \beta_1 * t_i \qquad \text{\# In model 2}$$

$$\log(\frac{P_{Vi}^{T}}{1-P_{Vi}^{T}}) = \alpha_i + \theta_i * \beta_1^{t_i} \qquad \text{\# In model 3}$$

116

, where $\mathrm{P_{Vi}}^R$ and $\mathrm{P_{Vi}}^T$ are the proportion of vaccinated individuals in the reference and target groups respectively, θ_i is the study-specific natural logarithm of the OR

We used a random effect model taking the between-study heterogeneity into account by assuming that θ_i were independent and sampled from a normal distribution centred around the mean log(OR) of carriage (μ) with a precision τ, such that $\theta_i \sim N(\mu, \tau)$ and $\tau = 1/\sigma^2$, where σ^2 is the between-study variance. A fixed effect was assumed for β_1.

Therefore, the vaccine efficacy at time t (VE$_{Ct}$) is as follows;

$$VEc_t = 1 - (e^{\mu} * t^{\beta_1}) \qquad\qquad \text{\# In model 1}$$

$$VEc_t = 1 - e^{(\mu + \beta_1 * t)} \qquad\qquad \text{\# In model 2}$$

$$VEc_t = 1 - e^{(\mu + \beta_1^{t})} \qquad\qquad \text{\# In model 3}$$

We used the same priors in all three models.

The models outputs were compared visually (Figure S2) as well as through the Deviance Information Criterion (DIC), with the smallest DIC suggesting the best model fit.

In the models of vaccine efficacy against carriage acquisition of all VT serotypes, the DIC was the same at 307.7, 307.4 and 307.0 for models 1, 2 and 3 respectively. Differences in DIC smaller than 5 are not considered meaningful in random effects meta-regression models.

Figure S2: Model 1 (left panel), model 2 (middle panel) and model 3 (right panel)
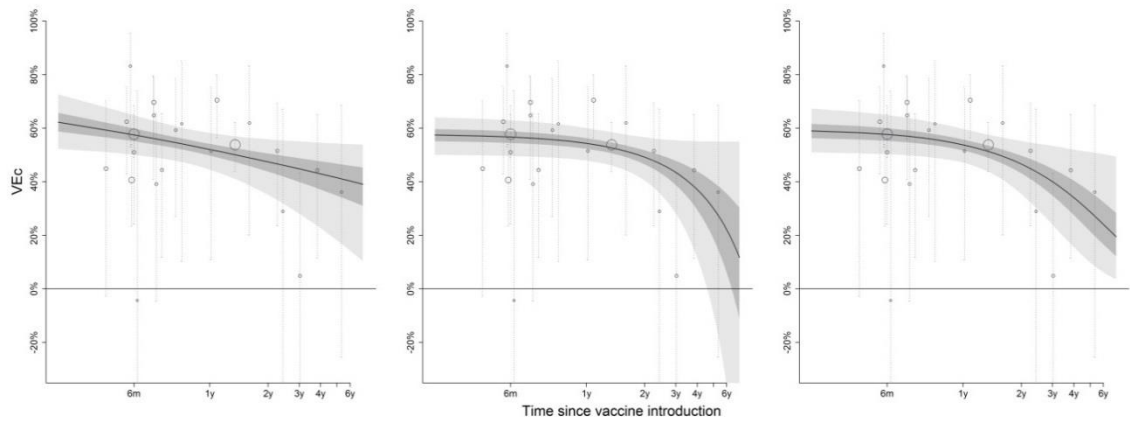
Legend: Left panel: model 1. Middle panel: Model 2. Right panel: Model 3. The plain line shows the model median, the dark grey shaded area the 50% credible interval (CrI) and the light grey shaded area the 95% CrI. The circles represent the point estimates of each individual study, with the size of the circle proportional to the study size, and the dotted vertical lines show the 95% confidence interval for each study

The DIC for the modelling of each individual serotype and each model considered are shown in Table S1.

The smallest DIC values were consistently seen for model 1 (the main model presented) – with the exception of serotype 9V –, but the difference in DIC values between models was not considered significant, except for 19F for which model 3 was outperformed by the two other models.

Hence, model 1 was presented as the main model in this paper based on a priori assumptions about the waning of vaccine efficacy, rather than on strong statistical grounds when comparing model 1 to the two other models.

Table S1: Deviance Information Criterion (DIC) values for ST-specific models, comparing each of the three models considered

Serotype	Deviance Information Criterion for each model		
	MODEL 1	MODEL 2	MODEL3
4	116.7	117.0	119.0
6B	193.2	193.2	193.5
9V	142.9	142.0	143.5
14	174.4	175.3	178.4
18C	151.9	151.9	153.6
19F	193.4	193.4	199.5
23F	192.8	193.3	193.7
6A	192.7	192.8	193.5

4. Research paper 3: Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: A survey in southwest Uganda

O le Polain de Waroux[1], S Cohuet[2], D Ndazima[3], A J Kucharski[1], A Juan-Giner[2], S Flasche[1], E Tumwesigye[5], R. Arinaitwe[3], J Mwanga-Amumpaire[3,4], Y Boum II[3], F Nackers[6], F Checchi[7], R F Grais[2], W J Edmunds[1]

1. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom

2. Epicentre, Paris, France

3. Epicentre, Uganda Research Centre, Mbarara, Uganda

4. Mbarara University Of Science and Technology (MUST), Mbarara, Uganda

5. Kabwohe Medical Research Centre, Kabwohe, Uganda

6. Epicentre, Brussels, Belgium

7. Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom

LONDON
SCHOOL *of*
HYGIENE
&TROPICAL
MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Olivier le Polain de Waroux
Principal Supervisor	W John Edmunds
Thesis Title	Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	NA		
When was the work published?	NA		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Nature Scientific reports
Please list the paper's authors in the intended authorship order:	O le Polain de Waroux, S Cohuet, D Ndazima, A J Kucharski, A Juan-Giner, S Flasche, E Tumwesigye, R. Arinaitwe, J Mwanga-Amumpaire, Y Boum II, F Nackers, F Checchi, R F Grais, W J Edmunds
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate designed the study, developed the protocol, developed and piloted study materials, trained study teams, helped supervise data collection, conducted the analysis and drafted the manuscript. Co-

	the study, the study design, protocol, the development and piloting of the study material, supervision and coordination of data collection in the field, provided input into the analysis and the manuscript. All authors critically revised the final manuscript and approved its final version

Student Signature: _____ **Date:** 18 November 2016

Supervisor Signature: W. J Edmonds **Date:** 18 November 2016

Abstract

Background

Quantification of human interactions relevant to infectious disease transmission through social contact is central to predict disease dynamics, yet data from low-resource settings remain scarce.

Methods

We undertook a social contact survey in rural Uganda, whereby participants were asked to recall details about the frequency, type, and socio-demographic characteristics of any conversational encounter that lasted for ≥5 minutes (henceforth defined as 'contacts') during the previous day. An estimate of the number of 'casual contacts' (i.e. <5 minutes) was also obtained.

Results

A total of 568 individuals were included. On average participants reported having routine contact with 7.2 individuals (range 1-25). Children aged 5-14 years had the highest frequency of contacts and the elderly (≥65 years) the fewest ($P<0.001$). A strong age-assortative pattern was seen, particularly outside the household and increasingly so for contacts occurring further away from home. Adults aged 25-64 years tended to travel more often and further than others, and males travelled more frequently than females.

Conclusion

Our study provides detailed information on contact patterns and their spatial characteristics in an African setting. It therefore fills an important knowledge gap that will help more accurately predict transmission dynamics and the impact of control strategies in such areas.

4.1.   Introduction

Quantification of human interactions relevant to the spread of infectious diseases spread by close contact is essential to accurately predict their infection dynamics and the impact of control strategies (1, 2).

Detailed surveys of social mixing patterns have now been undertaken in a number of settings (2-12). Studies have shown that people tend to mix with other individuals of their own age (i.e. assortative mixing); however, the frequency of contact, the degree of intergenerational mixing and the characteristics of mixing tend to vary between settings, depending on factors such as household size, population density and local activities, among others (3-11).

Data from low-resource settings remain scarce, with only two studies in Africa published to date (10, 12). With the exception of a recent study from China (11), the spatial dispersal of social contacts relevant for transmission has often been overlooked, and there is – to our knowledge – no published information from low-income settings on the spatial characteristics of social contacts. Spatial mobility is particularly important for epidemic risk prediction of novel and re-emergent diseases, and for the optimization of routine control programmes (13).

To address this knowledge gap, we set up a study of social contacts relevant to the spread of infections transmitted through the respiratory route or by close contact, in rural southwest Uganda.

4.2.   Materials and methods

The study was conducted in four sub-counties of Sheema North Sub-District (southwest Uganda), an area with a total of about 80,000 inhabitants. About half (49%) of the district's population is <15 years. The area is primarily rural.

4.2.1. Study design

Between January and March 2014 survey teams undertook interviews of a subset of individuals who were also included in a survey of *Streptococcus pneumoniae* carriage, asking about their social contacts in the 24 hours preceding the survey, including the frequency, type and duration of encounters.

Individuals were selected from 60 clusters randomly sampled from the 215 villages and two small towns in the sub-county, with an inclusion probability proportional to the size of the village or town. Within each cluster 11 or 12 households were randomly selected. A household was defined as a group of individuals living under the same roof and sharing the same kitchen on a daily basis. One individual from each household was randomly selected from a list of predefined age groups to sample from within each cluster. When nobody in the household was from that age group, either someone from another age group was selected providing that the quota for that age group had not been reached in the cluster, or the closest neighbouring household was visited instead. In case of non-response, another attempt was made later in the day or the following Saturday. Our target sample size was 687, including all 327 individuals aged ≥15 years and a subsample of 90 children in each of the following age groups: <2 year olds, 2 – 4 years old, 5 – 9 years and 10 – 14 years old. Based on estimates from previous findings (12, 14), such sample size provided a precision of just over 1 contact on the mean number of contacts per day, and enabled detection of a 20% difference in the average number of daily contacts by age group.

4.2.2. Data collection

Informed consent was sought for individuals aged > 13 years, and consent was sought from a parent or guardian otherwise.

Participants were asked to recall information on the frequency, type and duration of social encounters from the time they woke up the day before the survey until when they woke up on the survey day (~ 24 hours).

We defined contacts as two-way conversational encounters lasting for ≥5 minutes. Participants were first asked to list all the places they had visited in the previous 24 hours, the number of people they had contact with, their relationship with each individual mentioned, the age (or estimated age) of each listed contact and how long the encounter lasted for. Contacts involving skin-to-skin touch or sharing utensils passed directly from mouth-to-mouth were defined as 'physical' contacts. The questionnaire can be found in the Supporting Information.

We defined short contacts lasting less than 5 minutes as 'casual contacts'. Participants were only asked to estimate the number of casual contacts they had, based on pre-defined categories (<10, 10-19, 20-29, ≥30), but were not asked to provide detailed information about the nature of the encounter or the socio-demographic characteristics of the person met. Casual contacts are generally inaccurately reported in social contact surveys (7), particularly in a retrospective design, and most contacts important for the transmission of respiratory infections are believed to be close and non-casual rather than casual (6).

The questionnaire was designed in English, translated to Ruyankole, the local language, and back-translated to English for consistency. For children <5 years, parents were asked about their child's encounters and whereabouts. Children aged 5 – 14 years were interviewed directly, using a questionnaire with a slightly adapted wording to that used for adults.

The questionnaires and other study material can be found in the Supporting Information. For the conciseness of this thesis, not all study material (e.g. questionnaires for all age groups) are

appended to this Research Paper; however these are available on request and will be published with the manuscript.

Geographical coordinates from each participant's household and the centre of each village were taken using handheld GPS devices. The spatial identification of each location in the area was done by the research team during the preparation phase of the study. Geo-referencing of each village, hamlet or town in the area was done using GIS imagery as well as by travelling to the different villages to collect that information using handheld GPS devices. Given that some villages had very similar names, interviewers carried with them a list of all of those (>300) with them, so as to avoid transcription errors.

Questionnaires completed in the field were double entered on a preformatted data entry tool by two data managers working independently. Data entry conflicts were identified automatically and resolved as the data entry progressed.

4.2.3. Ethics statement

Approval was obtained from the Ethical review boards of Médecins Sans Frontières (MSF), the Faculty of Medicine Research & Ethics Committee of the Mbarara University of Science and Technology (MUST), the Institutional Ethical Review Board of the MUST, the Uganda National Council for Science and Technology (UNCST) and the London School of Hygiene and Tropical Medicine (LSHTM).

4.2.4. Analysis

4.2.4.1. Characteristics of social contacts by time, person and place

We analysed the frequency distribution of contacts for a set of covariates, including age, sex, occupation, day of the week, distance travelled, and type of contact. Encounters reported with

the same individual in different settings counted as one contact only. Straight-line distances between the centre point of all villages and towns in the dataset were calculated, and these were then used to evaluate how far people travelled, based on the reported names of villages and town where each reported encounter took place, and their own village or town of residence.

We used negative binomial regression to estimate the ratio of the mean contacts as a function of the different covariates of interest. Negative binomial was preferred over Poisson regression given evidence of over-dispersion (variance > mean, and likelihood ratio significant (P<0.05) for the over-dispersion parameter). We considered variables associated with contact frequency at p<0.10 for multivariable analysis, and retained them in multivariable models if they resulted in a reduction of the Bayesian Information Criterion (BIC).

Next, we explored whether people reporting a high frequency of casual contacts (≥10 casual contacts) differed from those reporting fewer contacts with regards to their socio-demographic characteristics. We did so using log-binomial regression to compute crude and adjusted relative risks (RRs) for having a high frequency In all analyses we accounted for possible within-cluster correlation by using linearized based variance estimators (15). Analyses were also weighted for the unequal probabilities of sampling selection by age group.

4.2.4.2.   Age-specific social contact patterns

We analysed the age-specific contact patterns through matrices of the mean number of contacts between participants of age group j and individuals in age group i, adjusting for reciprocity, as in Melegaro et al. (6).

If x_{ij} denotes the total number of contacts in age group i reported by individuals in age groups j, the mean number of reported contacts (m_{ij}) is calculated as x_{ij}/p_j , where p_j is

the study population size of age group j. At the population level the frequency of contacts made between age groups should be equivalent such that $m_{ij}P_j = m_{ji}P_i$. The expected number of contacts between the two groups is therefore $C_{ij} = (m_{ij}P_j + m_{ji}P_i)/2$. Hence, the mean number of contacts corrected for reciprocity (m_{ij}^C) can be expressed as C_{ij}/P_j.

We tested the null hypothesis of proportionate mixing by computing the ratio of observed mixing patterns to that of expected mixing patterns if social contact occurred at random. Under the assumption of random mixing, the probability of encounter between age groups thus depends on the population distribution in each age group, and the contact matrix under this random mixing hypothesis was calculated based on the percentage of population in each age group. The ratio of observed over expected contacts was then computed, and confidence intervals were obtained through bootstrapping, with replacement, for a total of 1,000 iterations. This approach is similar to that by Read et al. in a study in China [11].

4.2.4.3. *Epidemic simulations*

Finally, in order to explore the infection transmission dynamics resulting from our contact pattern data, we simulated the spread of an immunizing respiratory infection transmitted through close contact in a totally susceptible population. The model contained nine mixing age groups, with a transmission rate at which individuals in age group j come into routine contact with individuals in age group i computed as $\beta_{ij} = m_{ij}^C / \omega_i$, where ω_i is the proportion of individuals in age group i, and qm_{ij}^C is the next generation matrix, with q representing the probability of successful transmission per contact event (16). We assumed q to be homogeneous and constant across all age groups and conducted a set of simulations for fixed values of q between 25% and 40%, in line with what has been reported with influenza

pandemic strains (16, 17). The basic reproduction number (R_0) – which corresponds to the average number of people infected by one infectious individual in a totally susceptible population – was calculated as the dominant eigenvalue of the next generation matrix. We took uncertainty estimates in the contact matrices (and hence final size outputs) into account by iterating the model on bootstrapped matrices.

We then computed the final epidemic size (i.e. the number of individuals who would have been infected during the epidemic) for each specific age group, based on a mass action model adapted to account for multiple age classes, as described in Kucharski et al. (18).

Estimates obtained using the contact data from Uganda were compared to that of Great Britain, using data from the POLYMOD study (4) for the latter and a similar approach to compute the mixing matrix. The model was parameterised with social contact data on physical contacts only, lasting ≥5 minutes, rather than all contacts, given that physical contacts generally seem to better capture contact structures relevant for the transmission of respiratory infections (6), and that the definition of physical contacts is more similar and comparable between studies than that of overall contacts.

All analyses were performed in Stata13.1 IC and R version 3.2.

4.3.   Results

4.3.1.   Study population

A total of 568 individuals participated in the survey. This corresponded to an overall response rate of 83%, higher among ≥15 years old (98%), and lower among under 2s (68%), 2-4 year olds (64%), 5 – 9y olds (82%) and 10 - 14y olds (57%). There were more female (58%) than male respondents, but this differed by age group, with fewer females in young age groups and more adult females than males (Table S2 in Supporting Information).

Table 6: Mean Number of Reported Contacts and Ratio of Means By Socio-demographic Characteristic of The Study Population, Sheema, Uganda, January – March 2014.

Variables	N	Mean number of contacts (95%CI)	Crude RoM (95%CI)	Age adjusted RoM (95%CI)
Age groups				
<2y	61	6.11 (5.42 ,6.81)	0.99 (0.83 ,1.17)	
2-4y	57	6.70 (6.08 ,6.81)	1.08 (0.95 ,1.23)	
5-9y	74	8.50 (7.79 ,9.20)	1.37 (1.18 ,1.59)	
10-14y	51	8.70 (7.52 ,9.90)	1.40 (1.19 ,1.66)	
15-24y	91	6.20 (5.51 ,6.88)	ref	
25-34y	55	6.89 (5.99 ,7.80)	1.11 (0.93 ,1.33)	
35-44y	54	7.73 (6.99 ,8.46)	1.25 (1.09 ,1.43)	
45-54y	46	7.74 (6.37 ,9.11)	1.25 (1.01 ,1.54)	
55-64y	26	6.27 (5.01 ,7.53)	1.01 (0.80 ,1.29)	
65+y	48	4.85 (4.18 ,5.53)	0.78 (0.64 ,0.95)	
Sex				
Female	328	7.05 (6.66 ,7.44)	ref	
Male	235	7.47 (6.82 ,8.11)	1.06 (0.96 ,1.18)	
Occupation/daily activity				
Pre-school child	93	7.00 (6.30 ,7.70)	1.06 (0.90 ,1.23)	1.26 (1.02 ,1.55)
Student	166	8.27 (7.56 ,8.98)	1.27 (1.10 ,1.46)	1.30 (1.05 ,1.63)
Office worker	4	11.34 (9.64 ,13.03)	1.81 (1.20 ,2.73)	1.70 (1.32 ,2.18)
Shop worker	34	6.82 (5.56 ,8.08)	1.03 (0.84 ,1.25)	1.03 (0.83 ,1.29)
Agriculture	105	7.30 (6.58 ,8.01)	1.12 (0.97 ,1.30)	1.12 (0.96 ,1.30)
Other manual worker	40	5.37 (4.36 ,6.38)	0.85 (0.70 ,1.04)	0.85 (0.68 ,1.06)
At home	60	6.43 (5.62 ,7.24)	ref	ref
Unemployed	11	6.44 (2.92 ,9.96)	0.83 (0.60 ,1.15)	1.22 (0.77 ,1.94)
Retired	8	4.77 (3.75 ,5.80)	0.71 (0.48 ,1.05)	0.89 (0.70 ,1.13)
Other/unreported	41	6.65 (5.80 ,7.51)	1.02 (0.85 ,1.23)	1.18 (0.98 ,1.43)
Day of the week				
Weekday	439	7.14 (6.79 ,7.49)	ref	
Sunday	124	7.50 (6.56 ,8.44)	1.05 (0.92 ,1.20)	1.04 (0.92 ,1.18)
Travel outside village/town in previous 24 hours				
No	427	6.57 (6.20 ,6.92)	ref	
Yes	139	9.04 (8.35 ,9.73)	1.38 (1.25 ,1.52)	1.35 (1.22 ,1.49)
Number of casual contacts				
<10	315	5 (1 -15)	Ref	Ref
10 -19	119	8 (2-23)	1.43 (1.28 ,1.59)	1.39 (1.25 ,1.55)
≥20	56	9 (2-25)	1.64 (1.44 ,1.86)	1.61 (1.43 ,1.83)
Don't know	76	8 (0-19)	1.52 (1.37 ,1.68)	1.45 (1.29 ,1.64)

CI= Confidence interval; RoM= ratio of means

131

The mean household size was 5.3 (median 5, range 1 – 18). Almost all (98%) school-aged children aged 6 – 14 years attended school or college. Among adults, agriculture was the main occupation and about 27% of the females were homemakers/housewives (Table 6).

4.3.1. Characteristics of contacts

4.3.1.1. Contacts (i.e. ≥5 minutes long)

A total of 3,965 contacts with different individuals were reported, corresponding to an average of 7.2 contacts per person (median 7, range 0 - 25) (Figure 10). The majority of contacts were physical, thus involving skin-to-skin contact (mean 5.1, median 5 (range 0 – 18)).

Figure 10: Number of Reported Contacts, Including All Contacts (A) and Physical contacts (B), Sheema, Uganda, January – March 2014.
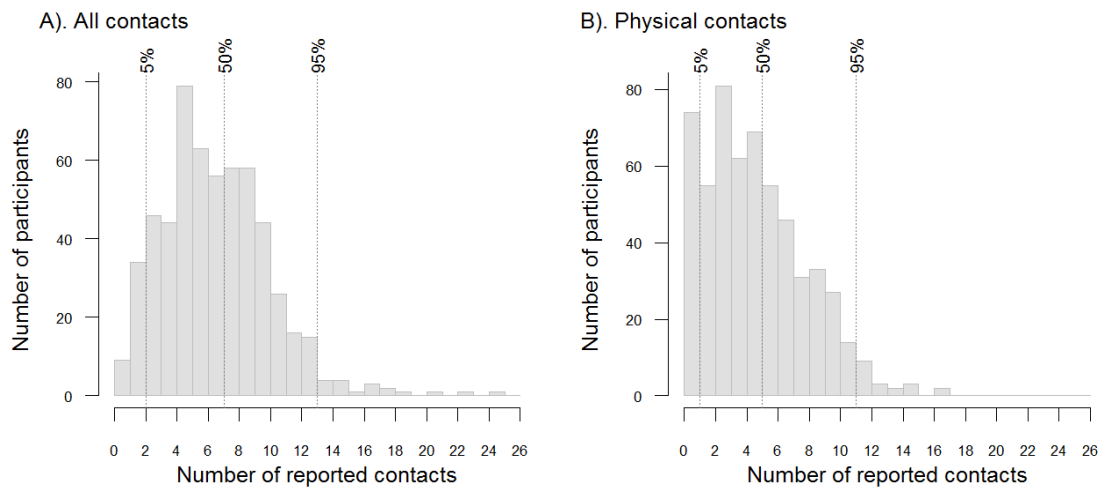
Legend: the vertical dotted lines represent the 5% centile, the median and 95% centile of the total number of reported contacts

Over half of all contacts (n= 2,060 (52%)) were with household members, 627 (16%) with other relatives, 873 (22%) with colleagues/friends/schoolmates and 402 (10%) with other individuals. The duration of routine contacts is shown in Figure S3 (Supporting Information).

Most contacts (82%) were with individuals who would be normally seen daily, 520 (13%) with people normally seen at least weekly, 4% with people met more rarely and 1% of the reported contacts were with people that the participants had never met before.

We found marked differences in the number of contacts by age group, but not by sex. School-aged children reported the highest daily number of contacts, while the elderly had the fewest (Table 6). There was no difference in the mean number of contacts for individuals living in the district towns of Kabwohe and Itendero (n=43) and the 523 others living in surrounding villages (2-sided P=0.79). Table 6 provides further details about the population characteristics, the mean number of contacts by socio-demographic and other covariates, as well as the ratio of mean contacts by covariate. Age was the only confounding factor.

Overall, contacts tended to be assortative, as shown by the strong diagonal feature on Figure 11, with most of the intergenerational mixing occurring within households (Figure 12). Only teenagers and adults reported non-physical contacts (Figure 12). The quantification of assortativity can be seen in Figures S4a-c, which show the ratios of observed contacts, as obtained in the survey but corrected for reciprocity, to that of expected contacts under the proportionality assumption, for all contacts and physical contacts only. The results show age-assortativity of contacts, for all age groups (other than <2 year olds) for all contacts, and primarily for school-aged children when considering physical contacts only.
Reciprocity correction accounted for the differential reporting between age groups, particularly higher frequency of contacts reported by small children with older age groups than older age groups reported (Figure S5).

There was no statistical difference in the average number of contacts between weekend (Sunday) and weekdays (Monday, Tuesday, Thursday and Friday) (Table 6). As shown in Table 6,

mean number of reported contacts on Sundays was 7.50 (95%CI 6.56; 8.44), slightly higher on average than on weekdays, where the average was 7.14 (6.79; 7.49), which was not statistically significant (P=0.229). The balance of respondents reporting contacts from weekdays and weekends reflected the normal proportion of week vs. weekend days in a normal week.

Taking into consideration that Saturday is sometimes only half a weekend day, the proportion of weekend days covered during the survey is representative of the proportion of weekend days in a week, whether the weekend is 1.5 days or 2 days long, as shown in the Figure S6.

About a quarter (n=136 (24%)) of participants reported social encounters outside their village of residence, and about 12% of contacts occurred outside participants' village of residence. The majority (56%) of people who travelled outside their village went to places located within a 5km radius from the centre point of their village of residence, and 90% stayed within 12km (Figure 13). Adult males tended to travel more than females (Figure 13). Overall, 29% of males had contact with someone outside their village, compared to 20% females (designed-based χ^2, P=0.0406). Most (87%) children under five years of age stayed in their village, whereas about a quarter or more individuals travelled outside their village among 5 – 14 years old (25%), 15-44 years old (32%) and ≥45 years of age (25%) , a difference by age group which was significantly different (P=0.0081).

When stratifying by age, the difference between sex were more marked, with no statistical difference between males and females <5 years of age (P=0.296) or 5 – 14 year olds (P=0.272), but marked differences among adults (≥15 years old), with 42% of males travelling outside of their village compared to 24% of females (P=0.0037).

Figure 11 : Average Number of Reported Contacts By Age Group, Sheema, Uganda, January – March 2014
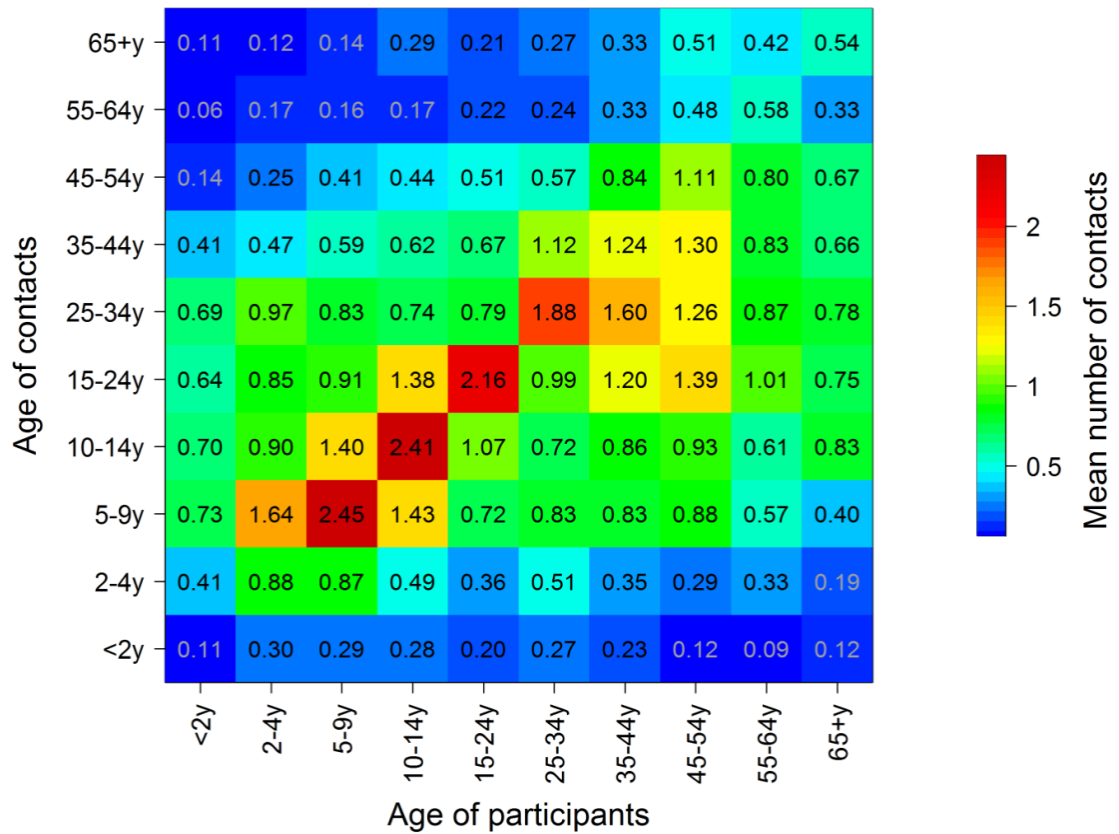
Legend: Numbers in each cell represent the average number of contacts between between age groups corrected for reciprocity, and 95% confidence intervals are shown in brackets

Most contacts made outside the household as well as those with individuals outside participants' village were mostly assortative (Figure 12), and the proportion of contacts outside the village was different by age group (P<0.001); higher among adults, increasingly so as distance from home increased (Figure 13).

Figure 12: Contact Matrices With Household members and Non-Household Members (Left upper and lower panel), for Physical and Non-Physical Contacts (Middle upper and lower panel), and for Contacts Made Within and Outside the Village (Right upper and lower panel), Sheema, Uganda, January – March 2014
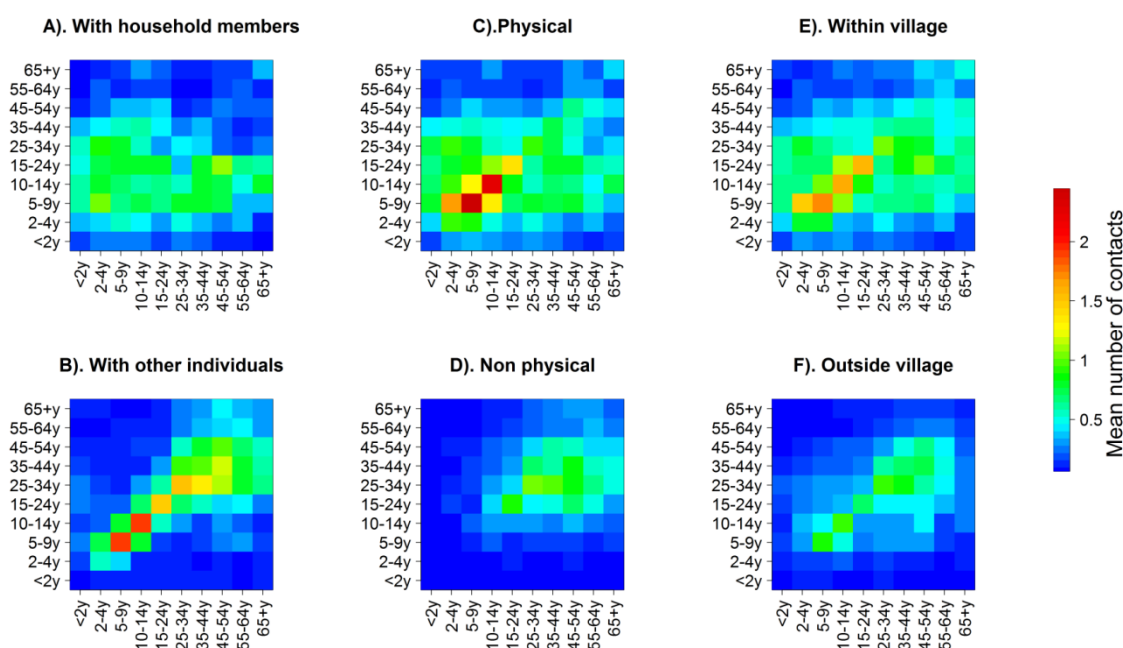
4.3.1.2. 'Casual' contacts (<5 minutes long)

Information on the number of casual contacts was reported by 490 (87%) participants. Among those, 64% (n=315) estimated they had fewer than 10 different contacts, 24% reported between 10 and 19 casual contacts, 6% reported between 20 – 29 contacts and 6% reported an estimated 30 contacts or more.

Individuals who reported high levels (i.e. ≥10 contacts) of casual contacts also tended to report more contacts (Table 6). We found no difference between those reporting high number of casual contacts (≥10) and others, by age, sex or day of the week (Table S3 in the Supporting Information). However, people whose primary activity was at home tended to reported fewer casual contacts than others, and there were about 50% more individuals reporting high levels of casual contacts among those who travelled outside their village.

Figure 13: Distance Travelled By Study Participants in the 24 Hours Preceding the Survey, Overall (A) and By Categories of Distance, Age and Sex (B), Sheema, Uganda, January – March 2014

A). Maximum distance (straight line) participants travelled in the previous 24 hours
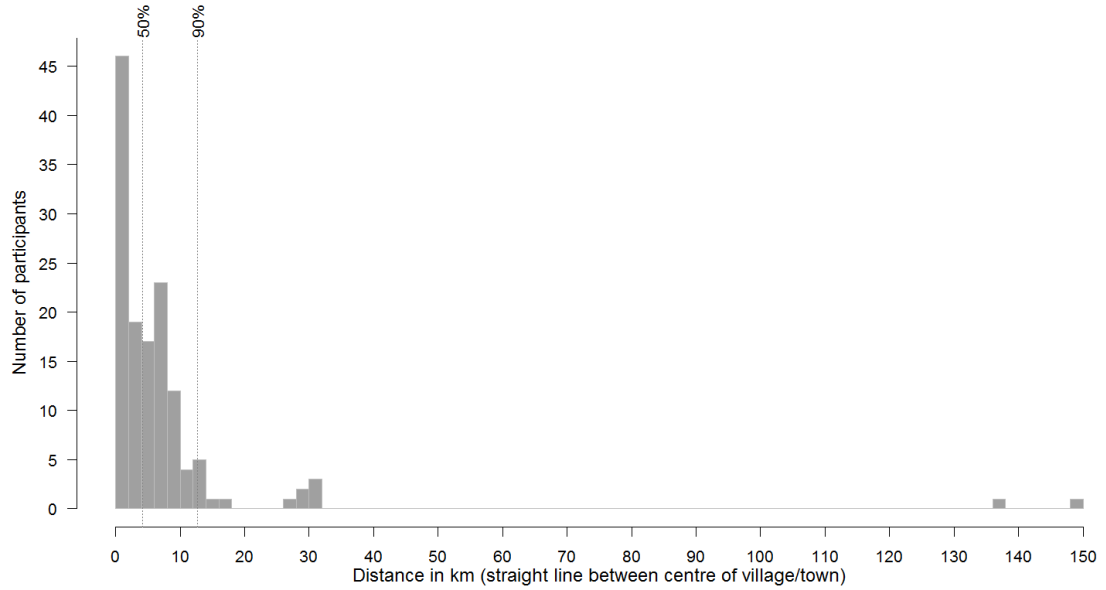
B). Categories by age and sex
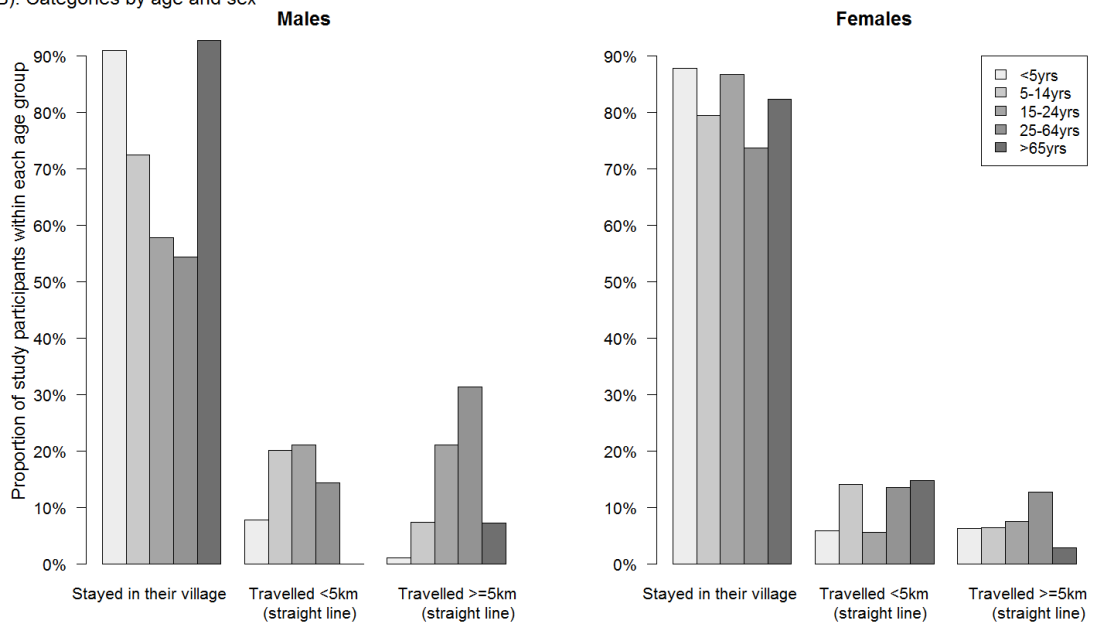
4.3.2. Epidemic simulations

Finally, we compared patterns of reported physical contacts in Uganda and Great Britain, and explored differences in the relative and absolute epidemic size by age group, as well as the corresponding R_0, for a hypothetical respiratory infection in an immune-naive population.

The number of reported physical contacts was similar between Uganda and Great Britain, with the average number of contacts by age group ranging from 3.2 (≥65 year olds) to 7.3 (2 − 4 year olds) in Uganda and from 3.3 (55 − 64 year olds) to 7.3 (10 − 14 year olds) in Great Britain. However contacts were more assortative in Britain than in Uganda (Figures 14A & B), some of which might be related to differences in household structures and number of household contacts, as contacts outside the household were mostly assortative (Figure 12).

The computed mean values of R_0 for a per contact infectivity value (q) ranging from 0.25 to 0.40 was slightly higher in Great Britain than in Uganda (1.51 to 2.41 vs. 1.40 to 2.24). Figure 14F shows the values for an infectivity parameter of 0.33. The proportion of people infected in younger age groups was also higher in Great Britain, and there were proportionally more adults infected in Uganda. However, given the differences in population structure, the total number of infections in the population was higher in Uganda than in Great Britain (Figures 14C − E).

Figure 14: Epidemic Simulation Using Matrices on Physical Contacts from Uganda (A) and Great Britain (B), for a Hypothetical Respiratory Infection In An Immune-Naïve Population, with the Proportion Infected by Age Group (C), the Epidemic Size by Age group (D), the Overall Proportion Infected (D) and the Basic Reproduction Number R0 (F).
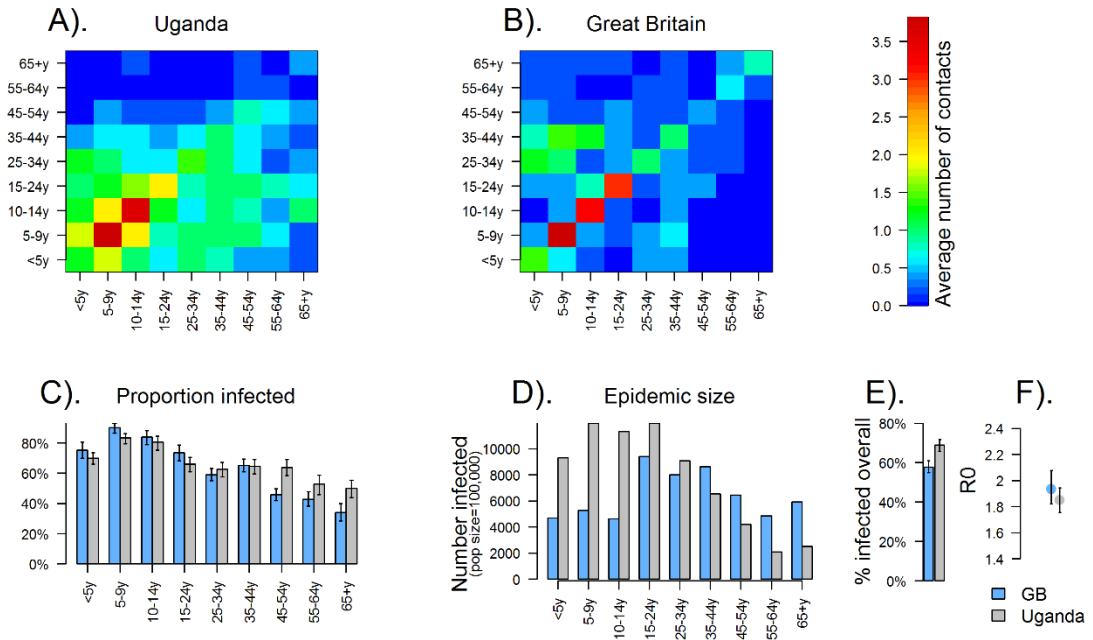
Legend: A: Matrix for physical contacts in Uganda. B: Matrix of physical contacts in Great Britain. C: Epidemic final size simulation: Proportion of individuals infected by age group in Great Britain (blue) and Uganda (grey), with error bars representing the 95% confidence interval. The results are presented for a q value of 33%. D: Epidemic size by age group, based on a total population size of 100,000 in Great Britain and in Uganda. E: Total proportion of people who were infected at the end of the epidemic in each setting. F: Estimates of R_0 for each setting, based on a q value of 33%, with dots showing the mean value and the bars showing the 95%

4.4.   Discussion

To our knowledge this is only the third study of its kind in Africa (10, 12), and the first one to specifically explore spatial patterns of social contacts. The quantification of mixing patterns is key to accurately model transmission dynamics and inform infectious disease control strategies (4). Having such data thus fills an important gap, particularly given the high burden of respiratory infections in low income settings (19, 20), and the risk of emerging and re-emerging diseases transmitted by close interpersonal contact, such as influenza (21), measles (22) or meningococcal meningitis (23).

Our findings share similarities with studies from Africa (10, 12) and other low or lower-middle income settings (14, 24), including the high contact frequency among school-aged children and that most contacts tend to be age-assortative. We also found substantial mixing between age groups, largely driven by intra-household mixing. This may result in a higher force of infection from children to adults than would be seen in high-income settings such as Great Britain, as our final size epidemic model suggests. The final size model should be seen as an illustration of how different social mixing patterns impact on disease epidemiology in different settings, rather than a specific quantification of the differences. It shows the importance of using setting-specific data when modelling disease dynamics and evaluate control strategies. Our data could be best applied to evaluate transmission dynamics and the impact of interventions for endemic diseases and current epidemics in non-naïve population, such as the ongoing large measles outbreak in neighbouring Democratic Republic of the Congo (22). In our final size model, it is also likely that our retrospective design resulted in underreporting compared to a prospective diary-based approach (25), which hampers comparisons between countries. In sensitivity analyses we explored the impact of potential underreporting in our retrospective survey design compared to a prospective diary-based approach (25), assuming a 25% under-ascertainment compared to a diary-based study, with homogeneous underreporting across age groups. In

140

such scenario, the proportion of infections across all age groups is predicted to be higher in Uganda than in Britain, disproportionally so in adults, and the R_0 to be higher too (see Figure S7 (Supporting Information)).

Our results also provide important insights into the local spatial dynamics of routine daily human interactions, showing that most contacts tend to occur within the vicinity of people's area of residence, that working age adult males travel most and young children and the elderly the least, and that contacts tend to be increasingly age assortative as people travel further away from home. Similar patterns were observed in rural and semi-urban China (11). Such findings have important implications to predict outbreak dynamics and control strategies given that interconnectedness between geographic patches is an essential factor driving epidemic extinction or persistence of epidemics hotspots and the effectiveness of control strategies. Studies of measles in Niger suggest that dynamics differ from that observed in high-income countries in the pre-vaccination era, likely due to different mixing patterns and weaker spatial connectivity (26, 27). This, together with important variations in vaccination coverage between local geographic patches (28-30), strengthens the need to account for spatial mobility when designing efficient control strategies in those settings. Optimal targeted interventions tailored to specific geographic clusters of high transmission have also been key considerations in recent cholera outbreaks in Africa, given the limited available vaccine doses (31, 32). Spatially targeted approaches are also central to outbreak control in the recent West African Ebola epidemic (33), and the current measles epidemic in the Democratic Republic of the Congo, which is sustained in part due to inadequate coverage of populations in less accessible geographical clusters (22, 34).

In our study the frequency of contacts was about half that of the number of contacts reported in Kenya, (10) or South Africa (12). Although differences between settings are expected, some of these are likely to be due to the exclusion of 'casual contacts' from our contact count. There

might be further differences linked to the definition of social contacts, which was based on conversational encounters in our study but not in the Kenyan study (10). When defining contacts based on conversational exchanges the household setting tends to dominate over other settings, compared to a more inclusive definition (8).

Both our contact definition and the retrospective study design may have resulted in more stable, regular contacts being reported over others. However, the extent to which a more inclusive definition reflects contact events relevant for transmission remains unclear. Modelling studies suggest that close interpersonal rather than short casual contacts matter more for transmission of respiratory infections (6). In addition, for modelling purposes the age-specific structure of relative contact frequency matters more than the actual reported frequency, as matrices are scaled to fit epidemiological data. Our retrospective interview-based design thus offers a simpler and easier alternative to prospective diary based approaches, particularly in such settings. Further research should explore what contact information is most relevant and how such data should best be captured. Research Paper 4 offers some insight into this question, by exploring contact types associated with pneumococcal carriage and acute respiratory symptoms based on contact and endpoint data (carriage and acute respiratory symptoms) collected at the same time from the same individuals.

Selection bias may have occurred to some extent, particularly given that more adult women were included than men. However, there was no significant difference in the number of contacts reported between males and females, including at the weekend, suggesting that selection bias was unlikely to be major. We also tried to reduce selection bias by interviewing on Saturdays people who were initially absent on the survey.

In conclusion, our study fills an important gap for two main reasons. First, we provide information by detailed age groups about social contacts and mixing patterns relevant to the spread of infectious diseases in a region where such data are scarce. Second, we also provide

some insights into spatial characteristics of social encounters. Although this has increasingly being recognized as an important component in evaluating epidemic risk and in the design of efficient control strategies, it has not previously been quantified in low-income settings, and should be explored further. Our study thus provides essential evidence to inform further research and infectious disease modelling work, particularly in similar rural African settings.

Conflict of interest

The authors declare they have no conflict of interest.

References

1.   Heesterbeek H, Anderson RM, Andreasen V, et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science* 2015;347(6227):aaa4339.

2.   Read JM, Edmunds WJ, Riley S, et al. Close encounters of the infectious kind: methods to measure social mixing behaviour. *Epidemiology and infection* 2012;140(12):2117-30.

3.   Edmunds WJ, O'Callaghan CJ, Nokes DJ. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proceedings Biological sciences / The Royal Society* 1997;264(1384):949-57.

4.   Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine* 2008;5(3):e74.

5.   Beutels P, Shkedy Z, Aerts M, et al. Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. *Epidemiology and infection* 2006;134(6):1158-66.

6.   Melegaro A, Jit M, Gay N, et al. What types of contacts are important for the spread of infections?: using contact survey data to explore European mixing patterns. *Epidemics* 2011;3(3-4):143-51.

7.   Smieszek T, Burri EU, Scherzinger R, et al. Collecting close-contact social mixing data with contact diaries: reporting errors and biases. *Epidemiology and infection* 2012;140(4):744-52.

8.   Bolton KJ, McCaw JM, Forbes K, et al. Influence of contact definitions in assessment of the relative importance of social settings in disease transmission risk. *PloS one* 2012;7(2):e30893.

9.   Kretzschmar M, Mikolajczyk RT. Contact profiles in eight European countries and implications for modelling the spread of airborne infectious diseases. *PloS one* 2009;4(6):e5931.

10.   Kiti MC, Kinyanjui TM, Koech DC, et al. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. *PloS one* 2014;9(8):e104786.

11.   Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proceedings Biological sciences / The Royal Society* 2014;281(1785):20140268.

12.   Johnstone-Robertson SP, Mark D, Morrow C, et al. Social mixing patterns within a South African township community: implications for respiratory disease transmission and control. *American journal of epidemiology* 2011;174(11):1246-55.

13.   Wesolowski A, Metcalf CJ, Eagle N, et al. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America* 2015;112(35):11114-9.

14.   Horby P, Pham QT, Hens N, et al. Social contact patterns in Vietnam and implications for the control of infectious diseases. *PloS one* 2011;6(2):e16965.

15.   Rogers WH. Regression standard errors in clustered samples. *Stata Technical Bulletin* 1993;13:19–23.

16.   Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American journal of epidemiology* 2006;164(10):936-44.

17.   House T, Inglis N, Ross JV, et al. Estimation of outbreak severity and transmissibility: Influenza A(H1N1)pdm09 in households. *BMC medicine* 2012;10:117.

18.   Kucharski AJ, Kwok KO, Wei VW, et al. The contribution of social behaviour to the transmission of influenza A in a human population. *PLoS pathogens* 2014;10(6):e1004206.

19.   Nair H, Simoes EA, Rudan I, et al. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. *Lancet* 2013;381(9875):1380-90.

20.   Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet* 2010;375(9730):1969-87.

21.   Radin JM, Katz MA, Tempia S, et al. Influenza surveillance in 15 countries in Africa, 2006-2010. *The Journal of infectious diseases* 2012;206 Suppl 1:S14-21.

22. Maurice J. Measles outbreak in DR Congo an "epidemic emergency". *Lancet* 2015;386(9997):943.

23. Burki T. Meningitis outbreak in Niger is an urgent warning. *The Lancet Infectious diseases* 2015;15(9):1011.

24. Grijalva CG, Goeyvaerts N, Verastegui H, et al. A household-based study of contact networks relevant for the spread of infectious diseases in the highlands of Peru. *PloS one* 2015;10(3):e0118457.

25. McCaw JM, Forbes K, Nathan PM, et al. Comparison of three methods for ascertainment of contact information relevant to respiratory pathogen transmission in encounter networks. *BMC infectious diseases* 2010;10:166.

26. Grais RF, Ferrari MJ, Dubray C, et al. Estimating transmission intensity for a measles epidemic in Niamey, Niger: lessons for intervention. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2006;100(9):867-73.

27. Ferrari MJ, Grais RF, Bharti N, et al. The dynamics of measles in sub-Saharan Africa. *Nature* 2008;451(7179):679-84.

28. Le Polain de Waroux O, Schellenberg JR, Manzi F, et al. Timeliness and completeness of vaccination and risk factors for low and late vaccine uptake in young children living in rural southern Tanzania. *International health* 2013;5(2):139-47.

29. Babirye JN, Engebretsen IM, Makumbi F, et al. Timeliness of childhood vaccinations in Kampala Uganda: a community-based cross-sectional study. *PloS one* 2012;7(4):e35432.

30. Satzke C, Turner P, Virolainen-Julkunen A, et al. Standard method for detecting upper respiratory carriage of Streptococcus pneumoniae: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* 2013;32(1):165-79.

31.   Azman AS, Luquero FJ, Rodrigues A, et al. Urban cholera transmission hotspots and their implications for reactive vaccination: evidence from Bissau city, Guinea bissau. *PLoS neglected tropical diseases* 2012;6(11):e1901.

32.   Luquero FJ, Banga CN, Remartinez D, et al. Cholera epidemic in Guinea-Bissau (2008): the importance of "place". *PloS one* 2011;6(5):e19005.

33.   Zinszer K, Morrison K, Anema A, et al. The velocity of Ebola spread in parts of west Africa. *The Lancet Infectious diseases* 2015;15(9):1005-7.

34.   Gil Cuesta J, Mukembe N, Valentiner-Branth P, et al. Measles vaccination coverage survey in moba, katanga, democratic republic of congo, 2013: need to adapt routine and mass vaccination campaigns to reach the unreached. *PLoS currents* 2015;7.

4.5. Supporting Information

4.5.1. Tables and Figures

Table S2: Age and Sex Distribution of Study Participants

Age category	Number (%) female	Number (%) male	Total number
<2 years	23 (38%)	38 (62%)	61
2 – 4 years	26 (45%)	31 (55%)	57
5 – 9 years	37 (50%)	37 (50%)	74
10 – 14 years	22 (43%)	29 (57%)	51
15 - 24 years	53 (58%)	38 (42%)	91
25 – 34 years	45 (79%)	12 (21%)	57
35 – 44 years	35 (64%)	20 (36%)	55
45 – 54 years	35 (76%)	11 (24%)	46
55 – 64 years	20 (77%)	6 (23%)	26
65+ years	34 (71%)	14 (29%)	48

Figure S3: The Reported Duration of Contact By Age Group Among Study Participants, Sheema District, January – March 2014
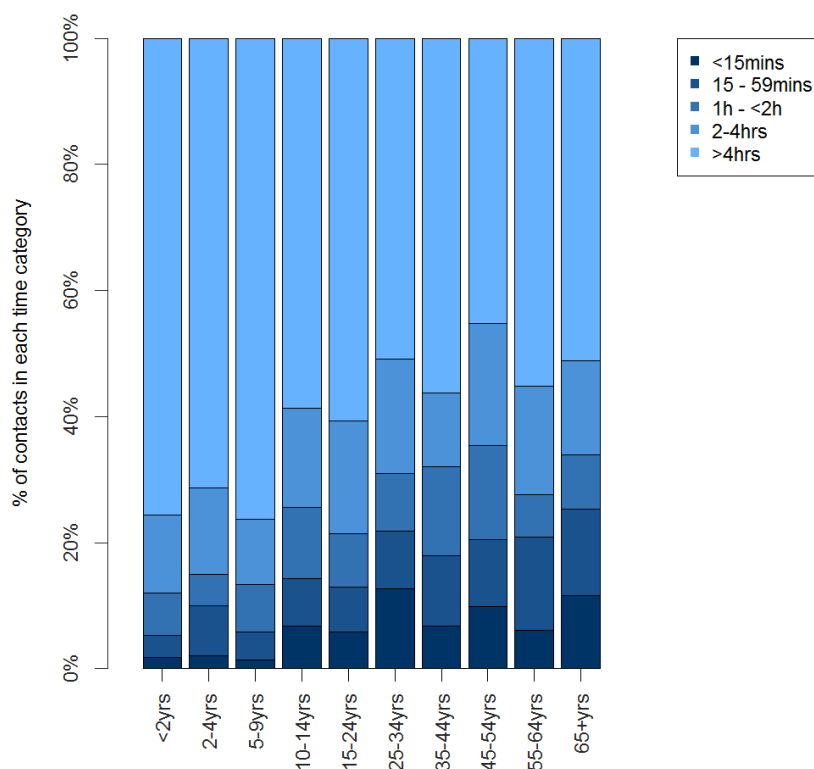
Table S3. Association Between Socio-demographic Variables and Level of Social Contacts

Variables	N	N (%) with high	Crude Risk Ratio	Adjusted RR (95%CI)
Age groups				
<2y	50	10 (20%)	0.57 (0.30 ,1.09)	
2-4y	47	14 (30%)	0.85 (0.50 ,1.45)	
5-9y	52	25 (48%)	1.37 (0.86 ,2.20)	
10-14y	43	19 (44%)	1.26 (0.80 ,1.99)	
15-24y	83	29 (35%)	ref	
25-34y	53	14 (26%)	0.76 (0.45 ,1.27)	
35-44y	49	27 (55%)	1.58 (1.11 ,2.25)	
45-54y	43	20 (47%)	1.33 (0.86 ,2.06)	
55-64y	23	7 (30%)	0.87 (0.42 ,1.80)	
65+y	47	10 (21%)	0.61 (0.32 ,1.14)	
Sex				
Female	297	98 (33%)	ref	
Male	193	77 (40%)	1.23 (0.96 ,1.59)	
Occupation/daily				
Pre-school child	81	22 (27%)	ref	ref
Student	132	53 (40%)	1.22 (0.80 ,1.86)	1.14 (0.74 ,1.74)
Office/Shop worker	34	19 (56%)	1.68 (1.11 ,2.54)	1.41 (0.90 ,2.21)
Agriculture/Manual	132	47 (36%)	1.07 (0.65 ,1.76)	1.00 (0.61 ,1.62)
At home	60	11 (18%)	0.55 (0.30 ,1.00)	0.51 (0.27 ,0.94)
Other	51	23 (45%)	1.38 (0.84 ,2.26)	1.27 (0.77 ,2.08)
Day of the week				
Weekday	385	138 (36%)	ref	
Sunday	105	37 (35%)	0.92 (0.67 ,1.26)	1.04 (0.92 ,1.18)
Travel outside				
No	374	118 (32%)	ref	
Yes	116	57 (49%)	1.58 (1.23 ,2.04)	1.54 (1.18 ,2.00)

Figure S 4a- c: Quantification of the Age-Assortativity of the Contact Matrices for All Contacts, Sheema District, January – March 2014. (S4A: median, S4B: lower 95% CI, S4C: upper 95%CI)
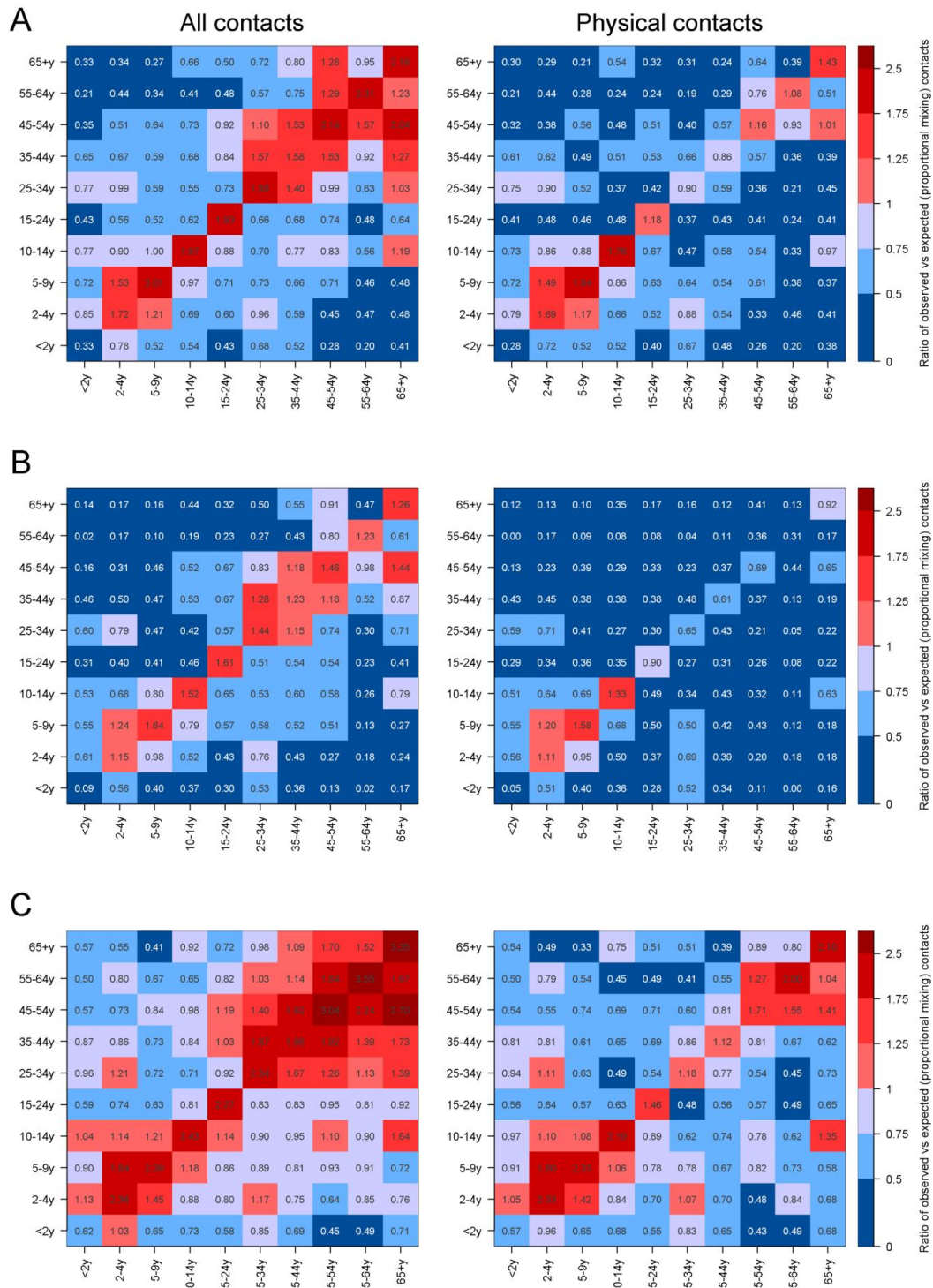
Legend: Left panel: all contacts; right panel: physical contacts. A: Median B:Lower 95%CI C:Upper 95%CI

Figure S5: Reciprocity Correction of the Contact Matrices for All Contacts, Sheema District, January – March 2014.
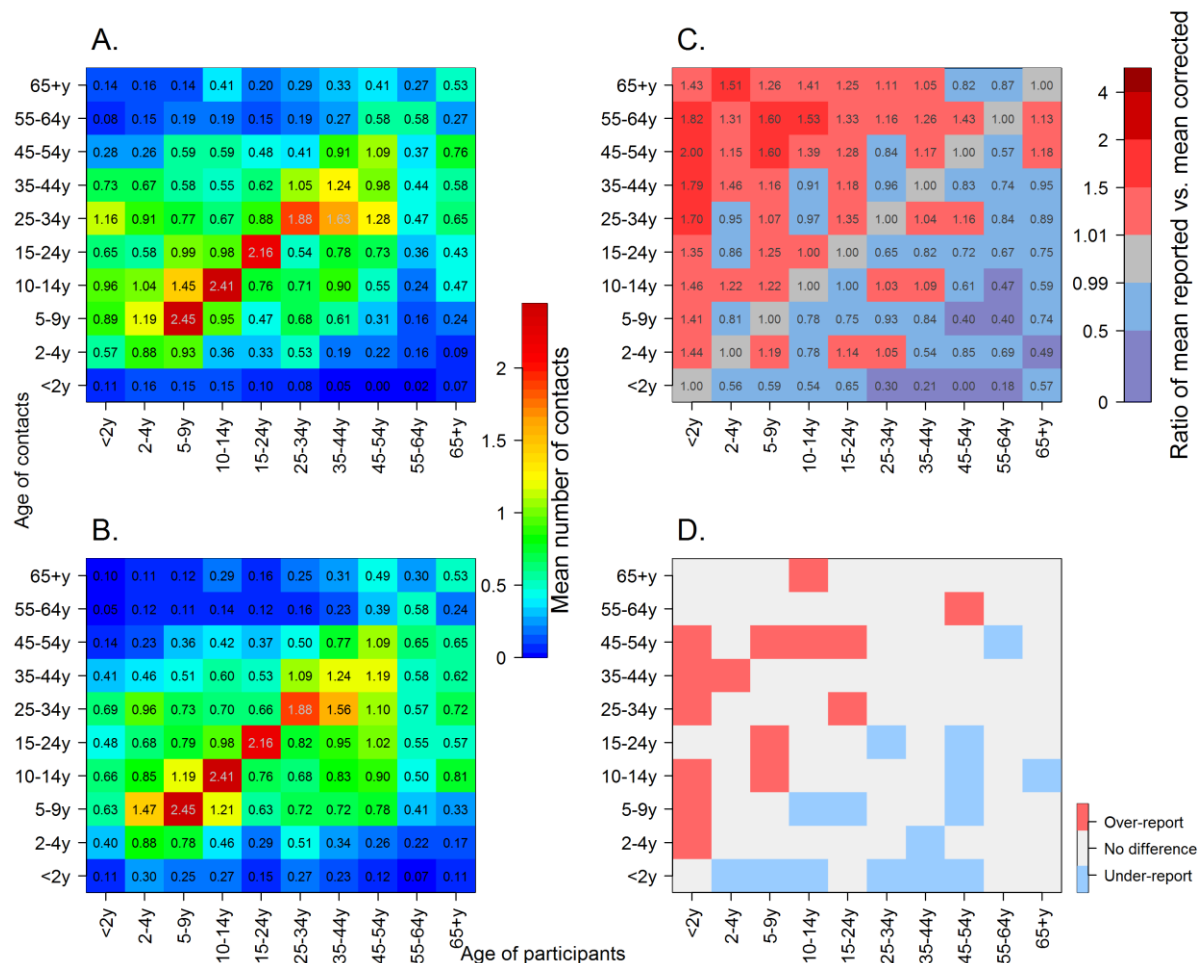
Legend: A) matrix for all reported contacts, not corrected. B) matrix for all reported contacts, corrected for reciprocity. C) Ratio of corrected over uncorrected matrices. Red cells illustrate where age-specific contacts were over-reported before correction, and blue cells under-reported. D) shows where participants significantly over-reported the number of age-specific contacts they had (upper 95% confidence bound) in red, significantly under-reported contacts in blue (lower 95% confidence bound), or where no significant adjustment was made (grey)

Figure S6: Proportion of contacts reported during weekend days, by age group.
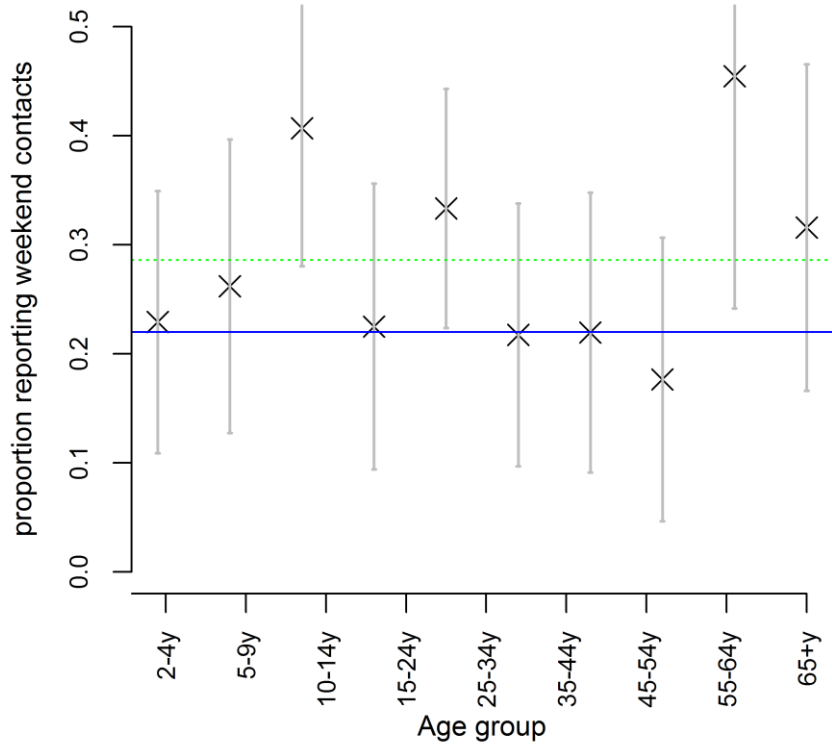
*proportion of weekend days in a week, if weekend = Sat pm + Sun(22%, blue line) and weekend =Sat all day+ Sun (29%, green line)

Figure S7: Epidemic Simulations Using Comparing Uganda and Great Britain, Assuming a 25% Underreporting Of Contacts In Uganda.
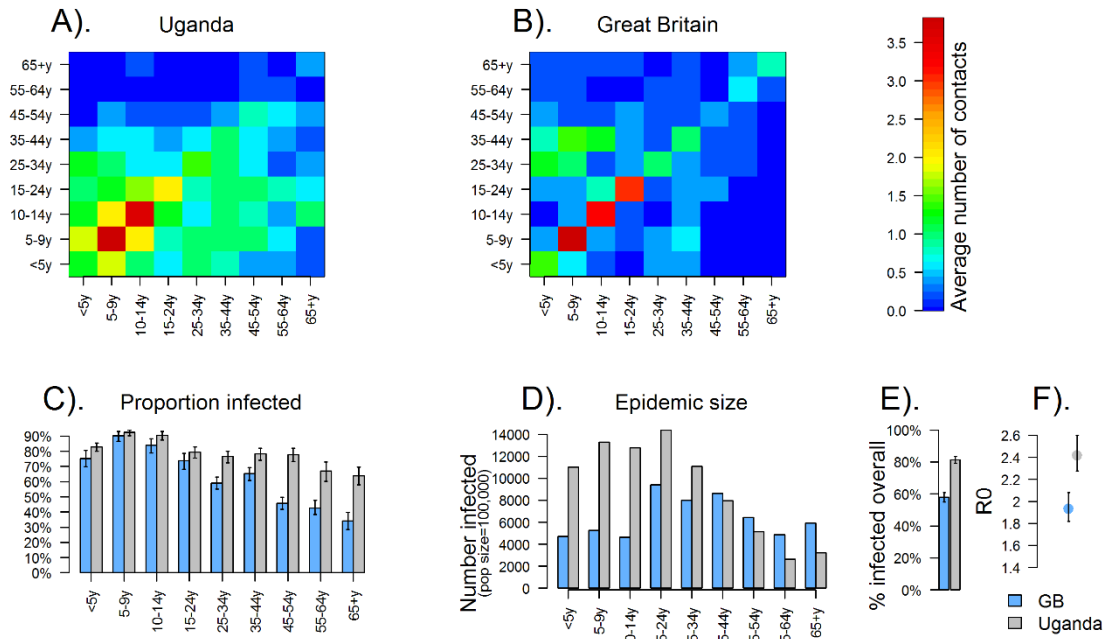
Legend: A) Physical Contacts from Uganda, B) Physical Contacts from Great Britain, C) Proportion Infected by Age Group, D) Epidemic Size by Age group, E) Overall Proportion Infected, and F) the Basic Reproduction Number R0.

4.5.2. Supporting File 1: Social contact questionnaire

There are three parts to this questionnaire. We will first ask a few questions about yourself/your child and the people you live with. Second, we will ask you questions about the different places you/your child went to yesterday. Finally, we will ask you to remember who you/your child met during the day in each of the places you attended, and how long for.

1. *[ID and dates]*

   1.1. *[Individual participant ID]:* |_|_|_|_|

   1.2. *[Interview date (dd/mm/yy)]*: |_|_|_|_|_|_|

   1.3. *[Date of the "surveyed day" (dd/mm/yy)]* : |_|_|_|_|_|_|

   1.4. *[Day of the week of the "surveyed day"]*
   1= Monday
   2= Tuesday
   3= Wednesday
   4= Thursday
   5= Friday
   6= Saturday
   7= Sunday

PART ONE: SOCIO-DEMOGRAPHIC CHARACTERISTICS

2. We would first like to ask a few questions about yourself/your child as well as about the people living with you/your child in your household.

   2.1. Age |_|_| years

   2.2. *[Sex of the participant]*
   1= Male

   2 = Female

   2.3. What is your/your child's primary occupation or daily activity *[note to the interviewer: this is the activity the participant spends the most time doing on a daily basis]?*
   1 = pre-school child;

   2= school/college/university student

   3= office worker

   4= shop worker

5= agriculture

6= manual worker

7= housewife

8= unemployed

9= retired

10= others

-1= don't know

-2 = refused

2.4. Including yourself/your child, how many people live in your household? **A household is the group of individuals living under the same roof and sharing the same kitchen on a daily basis**
|__|__| [-1 for don't know and -2 for refused]

2.5. How many bedrooms are in your household (only count living and sleeping rooms, exclude bathroom, kitchen)? |__|__| [-1 for don't know and -2 for refused]

3. We will now ask three questions about contacts you may have with animals
3.1. Do you/your child touch the following animals at least once per week?:
1= chickens, ducks, gees

2= cows, pigs, goats, sheep

3= rodent

4= primates

5= bats

6= antelope

7= none of the above

3.2. In the last month, have you/has your child been bitten or scratched or cut by any of these animals?
1= chickens, ducks, gees

2= cows, pigs, goats, sheep

3= rodent

4= primates

5= bats

6= antelope

7= none of the above

3.3. In the last month, have you/has your child killed, butchered or cooked any of these animals?

1= chickens, ducks, geese

2= cows, pigs, goats, sheep

3= none of the above

4. Including yourself/your child, could you please list all individuals in your household, as well as the family links between yourself/your child and each of the household members. **[define the household again]**

Household member initials/name	Household member ID			Sex		Age in years	Relationship								
				M=1	F=2		me	Spouse	Sibling	Child	Parent	Grandparent	Uncle/Aunt	Other family	Unrelated
	H	0	1	1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9
	H			1	2		1	2	3	4	5	6	7	8	9

PART TWO: SETTINGS

5. How far do you/your child travel(s) outside your village or town and how often?

Geographical setting	How often to you travel to those places outside your village or town?						How long do you spend in that place when you go				
	Most days of the weeks	At least once a week	At least once a month, but	Less than once a month	< once a month	Never	<1h	1-2h	Half a day	A whole day	Never
To a place in another village/town <5km away (<1 hour walk)	1	2	3	4	5	6	1	2	3	4	5
To a place in another village/town ≥5km away	1	2	3	4	5	6	1	2	3	4	5

6. In the last week,
   6.1. What is the name of the village/town of the furthest place from home you/your child went to? _____

   6.2. [Write the geographic ID of the abovementioned village/town **(see geographic ID form)**] |__|__|

7. Where did you/ your spend child time yesterday **[between the time of wake up yesterday and the time of wake up today]**

Setting ID			Type of place										If Other, where did you go? *[free text]*	Name of the village or town	Village/town with geographic ID		Time spent in that place				
			Home	Another house	work	School	Place of worship	Transport	Leisure	Shop	Garden	Other					<15mins	15mins – <1h	1 - <2h	2-<4h	>4h
S	0	1	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	2	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	3	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	4	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	5	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	6	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	7	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	8	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	0	9	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	0	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	1	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	2	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	3	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	4	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	5	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	1	6	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5
S	2	0	1	2	3	4	5	6	7	8	9	10					1	2	3	4	5

PART THREE: CONTACTS

We will now ask you to remember who you were/your child was in contact with yesterday **[between the time of wake up yesterday and the time of wake up today].**

8. How many people have you/your child seen **for a very short period of time (<5mins)** and with whom you/your child have **exchanged at least three words** yesterday? (e.g. saying hello to someone on your way, seeing someone in a shop, seeing a few children at school in the playground, etc)

> 1= 0 – 9 people
> 2= 10 – 19 people
> 3= 20 – 29 people
> 4= >30 people
> 5= don't know

9. Contacts

We would now like you to remember any person with whom you spent 5 minutes or more and with whom you exchanged at least three words in each of those settings. Those individuals will be defined as your contacts. For each of them, we would like to know if you had physical contact or non-physical contact. Nonphysical contact happens when you haven't touched the person. Physical contact includes hand shaking, sharing a bike, kissing, embracing, and also sharing a glass or other utensils passed directly from mouth to mouth.

9.1 We will first ask a few more questions about the people you were/your child was in contact with **at home** yesterday **for more than 5 minutes *[yesterday is defined as the period from wake up yesterdat to the moment you woke up this morning]***

Contact initials/name [will be removed from the questionnaire after the interview]	Contact ID			Place ID			Age (yrs)	Sex (M/F)		Type of contact		Link to the contact					How often do you have contact with this person in general?					Total time spent with the person in that particular place				
								M=1	F=2	Physical	Nonphysical	Household member	Other relative	colleague or schoolmate	Friend	Other	Daily or almost daily	At least once a week	At least once a month	Less than once a month	Never met before	5 - <15 mins	15 – <1h	1h - <2h	2h - <4h	>4h
	C	0	1	S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S	0	1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

9.2 We will now ask a few more questions about the people you were/your child was in contact with in the other places yesterday for more than 5 minutes

Contact initials/name	Contact number ID		Place ID			Age (yrs)	Sex (M/F)		Type of contact		Link to the contact					How often do you have contact with this person in general?					Total time spent with the person in that particular place				
							M=1	F=2	Close - physical	Close - nonphsycial	Household member	Other relative	colleague or schoolmate	Friend	Other	Daily or almost daily	At least once a week	At least once a month	Less than once a month	Never met before	5 - <15 mins	15 – <1h	1h - <2h	2h - <4h	>4h
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C		S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
C			S				1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Note that for the survey there was no limit to the number of contacts that could be entered as surveyors were able to add additional sheets as needed.

A slightly adapted questionnaire was used for children and parents.

4.5.3. Supporting File 2: Guidelines for the survey team

Note that for clarity here is a summary of the guidelines for the social contact survey part of the entire survey (which included both a nasopharyngeal carriage survey, with nasopharyngeal sampling procedures, collection and transport of specimens etc.)

The field coordinator will call the Village Health Teams to ensure their help for data collection during the day of the survey (which will take place 48h later) in the village. VHTs will be told that they will receive incentives to participate in the survey. They will be in charge to inform the heads of the selected households in the village/ cluster that the survey teams will visit their household two days later and that they should try to remember their social encounters during the 24 hours preceding the survey day. For households for which a telephone number is available, the head of household will be also directly contacted by the study coordinator to notify them of the survey team visit two days later.

The teamss will be divided in two groups of three and two teams (A + B + C and D+E) . The teams will go by motorbike to visit the selected villages of the day. Teams will visit 10 households per day in one same village, and D and E will visit 15 households per day. Fewer households need to be visited by teams A, B and C as they will visit villages located further away and/or more sparsely populated whereas teams D and E will visit villages that are easier to access and survey.

The list of households to be visited by village/cluster, including names of the head of households, will be available. **A household is defined as people who are sleeping under the same roof and sharing the same meal on a daily basis**. No matter the age of the household members or the relation between the members.

Selection of households

The nurse and the VHT will introduce the team to the household head or to another adult member of the family (i.e. anyone aged 18 years or older) in their absence. If either no adult or nobody is present in the household at the time of the visit, the team will move to the next household on their list and will come back to the earlier household later during the day.

If the head of the household or the participant refuses to participate in the survey, or if swabbing is contra-indicated, the household will be replaced by another household from a replacement list provided (**see list of households form**).

 At 6pm the activities in the field will stop. Teams will then assess the number of households that were not available to visit and feed this information back to the field coordinator . The field coordinator will then call the household heads of the households that were missed to inform them again about the survey and tell them that teams will come to visit them again on the

Saturday of the same week. If the head of the household tells the field coordinator that no one will be present on the Saturday, or if the head of the household cannot be reached by telephone, those households will be replaced from a random list(**see list of households form**) , and contact will be made by the field coordinator with those replaced households.

The household head will be carefully introduced to the survey purpose. Social benefit and confidentiality should be strongly emphasized (see **information participant form**). The verbal consent of the head of the household should be obtained. If the head of the household refuses to participate, teams should not insist on their participation. Teams should make a note about their refusal on the list of households (refusal yes/no) and continue with the next household on the list.

Selection of participants within households

- All individuals who are members of the selected households in the village/cluster will be eligible for study participation except if:
  1. the participant or their representative either refuses or is not able to provide informed consent to participate
  2. if they have head or facial injuries that would contraindicate NP swabbing.

- Only one person will be selected per household. To ensure that the sample size is reached for each age group, the age group of the individual to select will be predetermined using the "age group selection for cluster form" (see **age group selection for cluster form**). If several individuals of the same required age group are members of the selected household, the person to be included in the survey will be randomly chosen.

- If no one in the household is from the selected age group, the surveyors are entitled to include someone from another age group instead, so long as the sampling quota has not been reached for all other age groups from the list of households to include.

- If no one can be included, the household should be replaced by another household from the replacement list (**see list of households form**).

- Once the individual is selected within the household, the individual or the caregiver of a person less than 18 years old will be carefully introduced to the survey purpose by the nurse **(see information participant form).**

- Written consent from the individual has to be obtained (see **consent form over18 and consent formunder18**). If the selected individual refuses to participate, do not insist on their participation. If another household member in the same age group agrees participation, undertake the survey with them. Otherwise, move to the next household on the list.

- When the selected member of the household or their caregiver has given written consent to participate in the survey the interviewer will take the GPS coordinates of the household (see **SOP GPS coordinates**).

Contact survey questionnaire

1. General comments

- You (the interviewer) will have to complete the contact survey questionnaire for the individuals marked with a cross in the **age group selection for cluster form.**

- There are three versions of the questionnaire, each of which are adapted to particular age groups: <5y olds, 5-14y olds and ≥15y olds.

- Each team will carry one example of each questionnaire in Ruyankole. The interview will be carried out in Ruyankole, but questionnaires should be completed in English. The interviewer will conduct the interviews, and the study nurse will record the information on the English questionnaire.

- For children <5 years of age, parents/guardians will be interviewed on their behalf. Use the appropriate questionnaire for this. When asking questions about their child, replace 'your child' with the child's name

- Any information for the interviewer or any question that can be filled in directly be the interviewer without the help of the interviewee (e.g. date, sex of the interviewee) are shown in italics between squared brackets on the questionnaire

- For questions with multiple choices, please circle the appropriate response. In case of mistake, strike through the answer and circle the correct one, with a little 'V' sign next to it to notify the change, as shown here below

    2.4. Does your daily activity (listed above) involve >20 short (<5mins) contacts with individuals?
    ~~① Yes~~
    √ ② No
    -1= don't know

- Write the page number on the bottom right corner of each page of the questionnaire, as you go along

- For any numbers (e.g. participant's age, dates), fill all boxes in. In case of a single digit number when a two-digit space is provided, write a zero in the first box (as shown below)

2.1. How old are you? |0|7| years

- During the interview you will ask the study participants to list their household contacts and their social encounters in the various settings where they spent their time during the previous day. As the number of such contacts and the number of places visited during a 24 hour period varies between individuals, the questionnaire comprises of tables (each row corresponding to one social contact in a particular setting) which can be extended if needs be just by adding extra pages to the questionnaire. Do not forget to add page numbers to the extra pages inserted.

2. Part one: Socio-Demographic Characteristics

- You will first explain the purpose of the survey and define what is meant by social contacts (including close physical and nonphysical contact and casual contact). This will have to be repeated during the interview when asking about the type of social encounters over a 24 hour period.

- You will then write the individual ID on the top left corner of each page of the questionnaire. The individual ID will also be used to label household (HH) contacts and other social contacts, where the first 4 digits correspond to the individual ID and the following digits to the contact-specific ID.

- The "surveyed day" refers to the day for which people are being asked to recall their social encounters. For example, if the survey takes place on a Monday at 10 am, the participants will be asked to recall their contacts between morning and evening of previous day (Sunday). The 'surveyed day' will thus be the Sunday.

- The first part includes questions about the individual's daily activity, their household size and structure, and contact with animals. Please go through each question in the same order than on the questionnaire.

- Question 4 asks to list all household members and their age, sex and the relationship of the study participants with them (sibling, parent etc). It is advisable to list each household member's name in the first column on the left hand side of the table, to make household contact identification easier later in the questionnaire.

3. Part two: Settings

- First explain to the participants what is meant by the previous day: from wake up to sleep.

- You will be asked to enter the number of the village/town visited based on a predefined list of settings **(see geographic area coding form)**

- Promptpeople to remember all the settings they went to by starting in the morning and then moving through their day to the evening.

4. Part tree: Contacts

- This is the longest part of the questionnaire. For each of the settings mentioned by the participants, you will be asked to record the number of social contacts and their characteristics.

- First start by telling the participant how social contacts are defined

- For each setting,**start by writing the names or the initials of the contacts,** then go back to each of the contacts mentioned and ask details about the characteristics of the contact (including type of contact, duration etc).

- The first setting will always be the household. As some of the social encounters are likely to take place with the same individuals in different settings, for each individual mentioned in the first setting (i.e. the household), you will ask whether contacts with the listed individuals have also taken place in other settings. If yes, you will already note the names or initials of such individuals in the tables recording the contacts in those other settings. This will also ensure that there is only ONE ID PER INDIVIDUAL, even if encounters with the same person take place in different settings. Similarly, as you progress through the interview, for each new contact (e.g. contacts in setting 2 that are not in setting 1), also ask whether contact was made in any other setting (e.g. 3, 4, 5 etc). Please see an example here below

Contact initials/name [will be removed from the questionnaire after the interview]	Contact ID		Place ID	Age (yrs)	Sex (M/F)		Type of contact		Link to the contact						How often do you have contact with this person in general?					Total time spent with the person in that particular place				
					M=1	F=2	Physical	Nonphysical	Household member	Other relative	colleague or schoolmate	Friend	Other		Daily or almost daily	At least once a week	At least once a month	Less than once a month	Never met before	<15 mins	15 - <1h	1h - <2h	2h - <4h	>4h
MAX	C	0	1	S 0 1	32	①	2	①	2	①	2	3	4	5	①	2	3	4	5	1	2	3	4	⑤
RVKIA	C	0	2	S 0 1		1	②	①	2	①	2	3	4	5	①	2	3	4	5	1	2	3	4	⑤
JANET	C	0	3	S 0 1	45	1	②	①	2	①	2	3	4	5	①	2	3	4	5	1	2	3	4	⑤
	C			S 0 1		1	2	1	2	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

- If the age of individuals is unknown, write down a best guess, with the letter A (i.e. **A**pproximation) before the age, as displayed below

- For each new setting, the contact ID number of new contacts (ie. not previously encountered in previously listed settings) will one number above the ID number of the last new contact mentioned in the previous setting. See example below

Contact initials/name	Contact number ID	Place ID	Age (yrs)	Sex (M/F)		Type of contact		Link to the contact					How often do you have contact with this person in general?					Total time spent with the person in that particular place				
				M=1	F=2	Close - physical	Close - nonphsycial	Household member	Other relative	colleague or schoolmate	Friend	Other	Daily or almost daily	At least once a week	At least once a month	Less than once a month	Never met before	<15 mins	15 – <1h	1h - <2h	2h - <4h	>4h
MARY	C 0 4	S 0 2	6	1	②	1	②	1	②	3	4	5	1	②	3	4	5	1	2	③	4	5
MAX	C 0 1	S 0 2	32	①	2	①	2	1	2	3	4	5	①	2	3	4	5	1	②	3	4	5

5. Research paper 4: Identifying human encounters that shape the transmission of Streptococcus pneumoniae and other respiratory infections

Olivier le Polain de Waroux[1], Stefan Flasche[1], Adam J Kucharski[1], Celine Langendorf[2], Donny Ndazima[3], Juliet Mwanga-Amumpaire[3,4], Rebecca F Grais[2], Sandra Cohuet[5],

W John Edmunds[1]

1. Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT, London, United Kingdom

2. Department of Research, Epicentre, 8 Rue Saint-Sabin, 75011 Paris, France

3. Epicentre Mbarara Research Centre, PO Box 1956, Mbarara, Uganda

4. Mbarara University of Science and Technology, Mbarara University, PO Box 1410, Mbarara, Uganda

5. Department of Field Epidemiology, Epicentre, 8 Rue Saint-Sabin, 75011 Paris, France

LONDON
SCHOOL of
HYGIENE
&TROPICAL
MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Olivier le Polain de Waroux
Principal Supervisor	W John Edmunds
Thesis Title	Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	NA		
When was the work published?	NA		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Proceedings of the Royal Society B
Please list the paper's authors in the intended authorship order:	Olivier le Polain de Waroux, Stefan Flasche, Adam J Kucharski, Celine Langendorf, Donny Ndazima, Juliet Mwanga-Amumpaire, Rebecca F Grais, Sandra Cohuet, W John Edmunds
Stage of publication	Not yet submitted

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate designed the study, helped coordinate the survey in the field, conducted the analysis and drafted the manuscript. Co-authors contributed to the laboratory analyses, data entry, study coordination,

	provided input in the analysis and data interpretation. and piloting of the study material. All authors critically revised the final manuscript and approved its final version

Student Signature: _____   **Date:** 18 November 2016

Supervisor Signature: _____   **Date:** 18 November 2016

Abstract

Background

Although patterns of social contacts are believed to be an important determinant of infectious disease transmission, there is little empirical evidence to back this up. Indeed, few studies have attempted to link individuals' risk of respiratory infection with their *current* pattern of social contacts. Here we assess whether the frequency and duration of different types of social encounters are associated with current pneumococcal carriage and self-reported acute respiratory symptoms (ARS).

Methods and Findings

We conducted a survey in Uganda in 2014, in which 566 participants were asked about their daily social encounters and about symptoms of ARS in the last two weeks. A nasopharyngeal specimen was also taken from each participant.

We found that the frequency of physical (i.e. skin-to-skin), long ($\geq$1h) and household contacts – which capture some measure of close (i.e. relatively intimate) contact –, was higher among pneumococcal carriers than non-carriers, and among people with ARS compared to those without, irrespective of their age. With each additional physical encounter the age-adjusted risk of carriage and ARS increased by 6% (95%CI 2-9%) and 9% (1-18%) respectively. In contrast, the number of casual contacts (<5 minutes long) was not associated with either pneumococcal carriage or ARS. A detailed analysis by age of contacts showed that the number of close contacts with young children (<5 years) was particularly higher among older children and adult carriers than non-carriers, while the higher number of contacts among people with ARS was more homogeneous across contacts of all ages

Conclusions

Our findings provide key evidence that the frequency of close interpersonal contact is important for transmission of respiratory infections, but not that of casual contacts. Such results strengthen the evidence for public health measures based upon assumptions of what contacts are important for transmission, and are important to improve disease prevention and control efforts, as well as inform research on infectious disease dynamics.

5.1.   Introduction

The transmission of respiratory infections is likely to depend on the frequency and age structure of human social contacts, as well as other factors including pre-existing immunity from prior infection or vaccination [1-3]. To understand the dynamics of such infections, studies to quantify social mixing patterns have been conducted in various settings, under the assumption that self-reported encounters reflect transmission probabilities of pathogens transmitted through close contact [2, 4-11]. Combined with disease transmission models, these data are increasingly being used to inform infection control policies [12, 13].

There is evidence from population-based models that self-reported social mixing patterns can reproduce observed aggregated data for chickenpox [14], mumps [4], parvovirus [15], influenza [4, 16, 17]) and whooping cough [18]. Moreover, it has been suggested that age-stratified social mixing patterns can capture individual influenza risk, as measured by a four-fold rise in neutralization titres over the course of an epidemic [17, 18]. However, as yet no study has directly linked individual's contact patterns with their risk of infection at the same point in time.

To establish how social behaviour shapes individual-level infection, we explored whether the frequency and duration of different types of social encounters were associated with an individual's risk of respiratory infection, using nasopharyngeal (NP) carriage of *Streptococcus pneumoniae* (the pneumococcus) and self-reported acute respiratory symptoms (ARS) as endpoints.

Streptococus pneumoniae is one of the main causes of pneumonia and sepsis globally [19], disproportionally so in low-income settings [19-21]. Colonization of the nasopharynx is a precondition to disease, and the main source of human-to-human transmission. Given that most episodes of carriage remain asymptomatic, social behaviour is unlikely to change as a

result of carriage, making pneumococcal carriage a more suitable endpoint than symptomatic illness to explore the association between social behaviour and infection risk, given that people tend to limit their contacts during symptomatic illness [22]. In addition, as natural immunity to carriage is weak [23], pre-existing immunity is less likely to confound associations between disease and social contact patterns than in studies using immunizing infections as biological endpoint [16, 17].

Using individually matched data on carriage, ARS and contact patterns collected in a rural and semi-urban area of South West Uganda, we explored whether the frequency and duration of different types of social encounters were associated with an individual's risk of respiratory infection.

5.2. Methods

5.2.1. Data collection

The study was conducted in Sheema North Sub-District (Sheema district, South-West Uganda) between January and March 2014. Sixty clusters were randomly selected from the 215 villages and two small district towns (Kabwohe and Itendero) in the study area, proportionally to the population size of each village and town. In each cluster 11 or 12 individuals were randomly sampled from different households to both answer questions about their social contacts and their history of respiratory illness in the last two weeks, as well as having a nasopharyngeal swab taken. A household was defined as the group of individuals living under the same roof and sharing the same kitchen on a daily basis.

For the social contact questionnaire, participants were first asked to list all the individuals with whom they had a two-way conversational contact lasting for ≥5 minutes during a period of approximately 24 hours prior to the survey day (from wake up the previous day until wake up

on the survey day). Such encounters were defined as 'ordinary contacts'. For each reported ordinary contact, participants (or their parent/guardian) were asked to estimate the contact's age (or estimated age), how long the encounter lasted for and whether it involved skin-to-skin touch or utensils passed from mouth to mouth (either of those defining 'physical contacts'). For very short social encounters (<5 minutes), which were defined as 'casual contacts' (e.g. seeing someone on the way, encounter in a shop etc.), participants were asked to estimate the number of encounters based on pre-defined categories (<10 contacts, 10-19 contacts, 20-29 contacts, ≥30 contacts), but not to provide further details about each contact.

Next, participants were asked about respiratory symptoms experienced in the two weeks prior to the survey, including any of the following: cough, runny nose, sneezing, sore throat, difficulty breathing.

Finally, after the interview was completed, a nasopharyngeal swab was taken from each participant. NP samples were collected, transported and analysed as per WHO guidelines [24]. NP swabs (flocked nylon swabs, COPAN, Italy) were inoculated in a skim milk tryptone-glucose-glycerol (STGG) medium, transported in cool boxes and frozen at the research laboratory at -20°C within 8 hours of collection. Specimens were inoculated onto a selective agar plate of 5 mg/L gentamicin-Columbia agar with 5% sheep blood and incubated at 37°C in 5% CO_2 atmosphere overnight. Pneumococcal identification was based on optochin susceptibility testing of all alpha-hemolytic colonies and bile solubility testing in case of intermediate susceptibility to optochin.

5.2.2. Statistical analysis

We first performed descriptive analyses, with age-specific probability weights to account for different inclusion probabilities by age at the design stage, and adjusted for the clustering by village through the use of clustered 'sandwich' variance estimators to account for possible

correlation within each of the sixty clusters [25]. We explored social contact patterns within and between age groups through contact matrices. In such matrices we report the mean number of ordinary contacts of participants in age group j with contacts in age group i (m_{ij}), adjusted for reciprocity, as in Melegaro et al. [15].

We then analysed whether and how the frequency distribution of contacts was associated with pneumococcal carriage or self-reported ARS.

We modelled the effect of 'ordinary' contacts (defined as contacts ≥5 minutes long) on carriage or respiratory symptoms as a function of contact frequency, through a log-binomial model with a robust variance estimator, and inclusion probability weights by age group. We treated contacts as continuous variables, but assessed departure from linearity through likely ratio tests and model comparisons of Bayesian Information Criterion (BIC). In multivariable analysis, we considered for inclusion any covariate significantly associated with the outcome at P<0.10 in univariable analysis. Model improvement was considered for any decrease in the BIC. Further details are provided in the Supporting Information.

Next, we used the same analytical approach to assess whether carriage and ARS were associated with the level of self-reported 'casual' contacts (i.e. <5 minutes long).

Finally, we explored differences in the social mixing matrices by status of pneumococcal carriage and ARS. To do so, we computed the ratio of the mean number of reported contacts by participants j with contacts i among carriers compared to non-carriers (R_{ij}^{C}) or symptomatic compared to asymptomatic individuals (R_{ij}^{R}), such that $R_{ij}^{C} = \dfrac{m_{ij}^{C}}{m_{ij}^{N-C}}$ and

$R_{ij}^{R} = \dfrac{m_{ij}^{R}}{m_{ij}^{N-R}}$, where C = carriers, R = participants with ARS and N = the total number of participants. Estimates were not adjusted for reciprocity given that subpopulations were not closed (e.g. contacts of carriers may not be carriers), and that our aim was to compare reported number of contacts rather than calculate a contact matrix. The uncertainty in reported values

and ratios was obtained through resampling techniques, drawing random samples from each m_{ij}, with the number of draws equal to the study population in each age group j. Ratios and uncertainty around them was obtained from the ratio of bootstrapped matrices.

The same approach as described above was used to compute the differences in m_{ij} with

$$D_{ij}^{C} = m_{ij}^{C} - m_{ij}^{N-C} \text{ and } D_{ij}^{C} = m_{ij}^{R} - m_{ij}^{N-R}.$$

All analyses were performed in Stata 13.1 IC and R version 3.2.

5.2.3. Ethics

Approval was obtained from the Ethical review boards of the London School of Hygiene and Tropical Medicine, Médecins Sans Frontières, the Faculty of Medicine Research & Ethics Committee of the Mbarara University of Science and Technology, the Institutional Ethical Review Board of the MUST, and the Uganda National Council for Science and Technology.

5.3. Results

5.3.1. Study population

Of the 687 individuals initially targeted for inclusion, 568 (83%) individuals from unique households responded to the survey, with data on both social contacts and carriage or ARS available from 566 participants. Participant's age spanned across age groups and the sex distribution was reasonably balanced (58% female). The majority (98%) of children aged 6 – 15 years attended school or college. More than a third of all adults (36%) worked in agriculture and 22% were homemakers/housewife.

On average, people made seven *'ordinary'* contacts (defined as contacts ≥5 minutes long), ranging from 0 to 25, the majority of which were physical (i.e. 'skin-to-skin' ordinary contact or indirect physical contact through utensils passed from mouth-to-mouth). There was no evidence that the average number of contacts differed by weekday or between weekdays and

weekends (P=0.623). Children aged 5-9 years reported most contacts and children <5 years the fewest (Figure 15). The majority of contacts made by children were physical. The proportion of non-physical contacts, contacts outside the household and contacts of short (<1h) duration was higher among teenagers and adults than younger children.

The most intense mixing tended to be between individuals of the same age group (i.e. assortative mixing), but there was also substantial mixing between age groups. Contact from and with children aged <10 years involved proportionally more physical touch than contacts between older children and adults (Figure 16A and B). There was no difference in contact patterns by sex, for all types of contacts considered.

Four hundred and ninety (87%) participants estimated how many casual contacts (i.e. <5 minutes in duration) they had the day before the survey. Over a third (36%) of these reported 10 casual contacts or more, and 11% reported more than 20 casual contacts. Among the 13% who did not know how many casual contacts they had, there were proportionally more children with over half (56%) under 10 years.

Figure 15: Number of reported contacts by age group and type of contact.
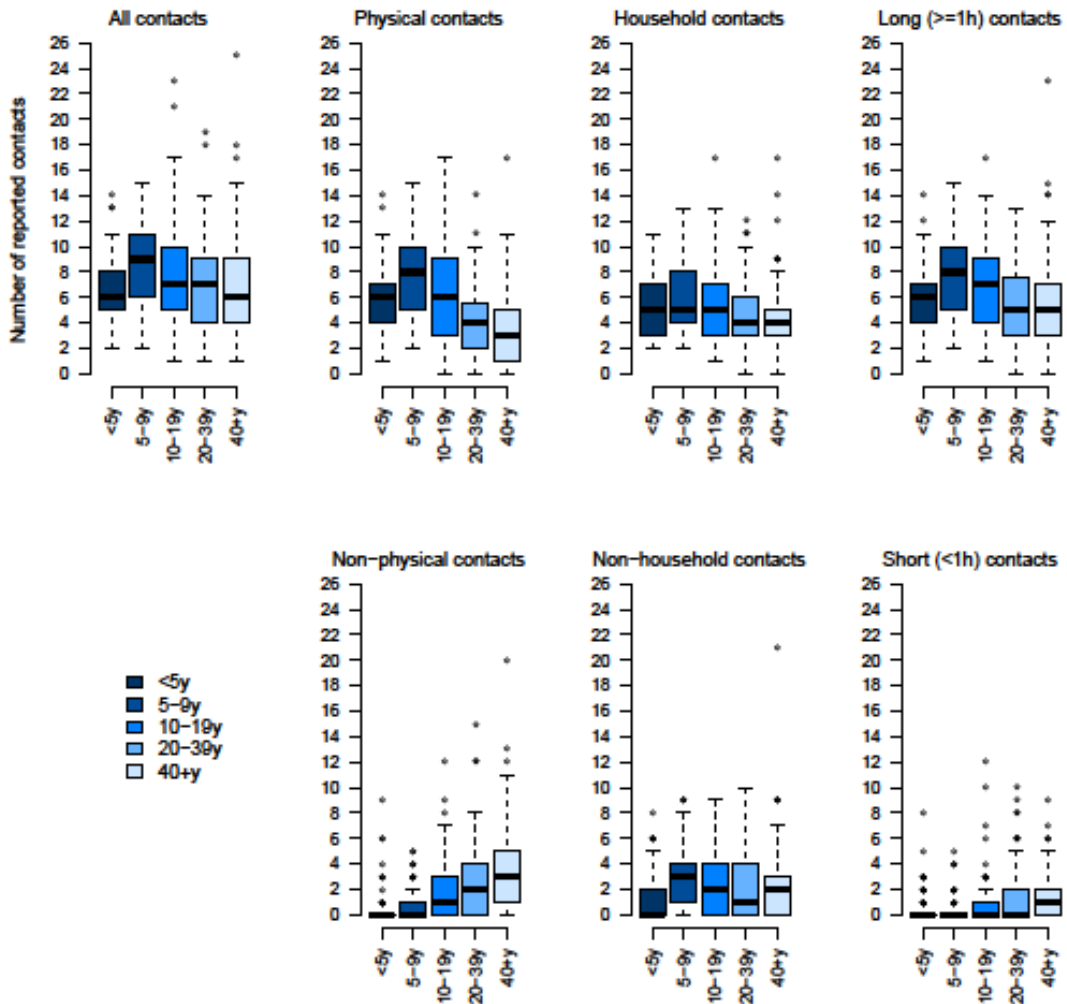
Legend: Boxplots showing the median (horizontal line), interquartile range (boxes) and 95% confidence interval around the mean (whiskers). The outliers are shown as bubbles above or under the whiskers

Four hundred and ninety (87%) participants estimated how many casual contacts (i.e. <5 minutes in duration) they had the day before the survey. Over a third (36%) of these reported 10 casual contacts or more, and 11% reported more than 20 casual contacts. Among the 13% who did not know how many casual contacts they had, there were proportionally more children with over half (56%) under 10 years.   The mean number of reported 'ordinary' contacts was higher among individuals reporting ≥10 casual contacts than those reporting fewer than 10

contacts (mean 8.9 vs 5.9, P<0.001), as well as among the 78 individuals who did not know how

many casual contacts they had (mean 8.8, P<0.001).

Figure 16: Contact matrices and prevalence of *Streptococcus pneumoniae* carriage and Acute
Respiratory Symptoms (ARS), by age.
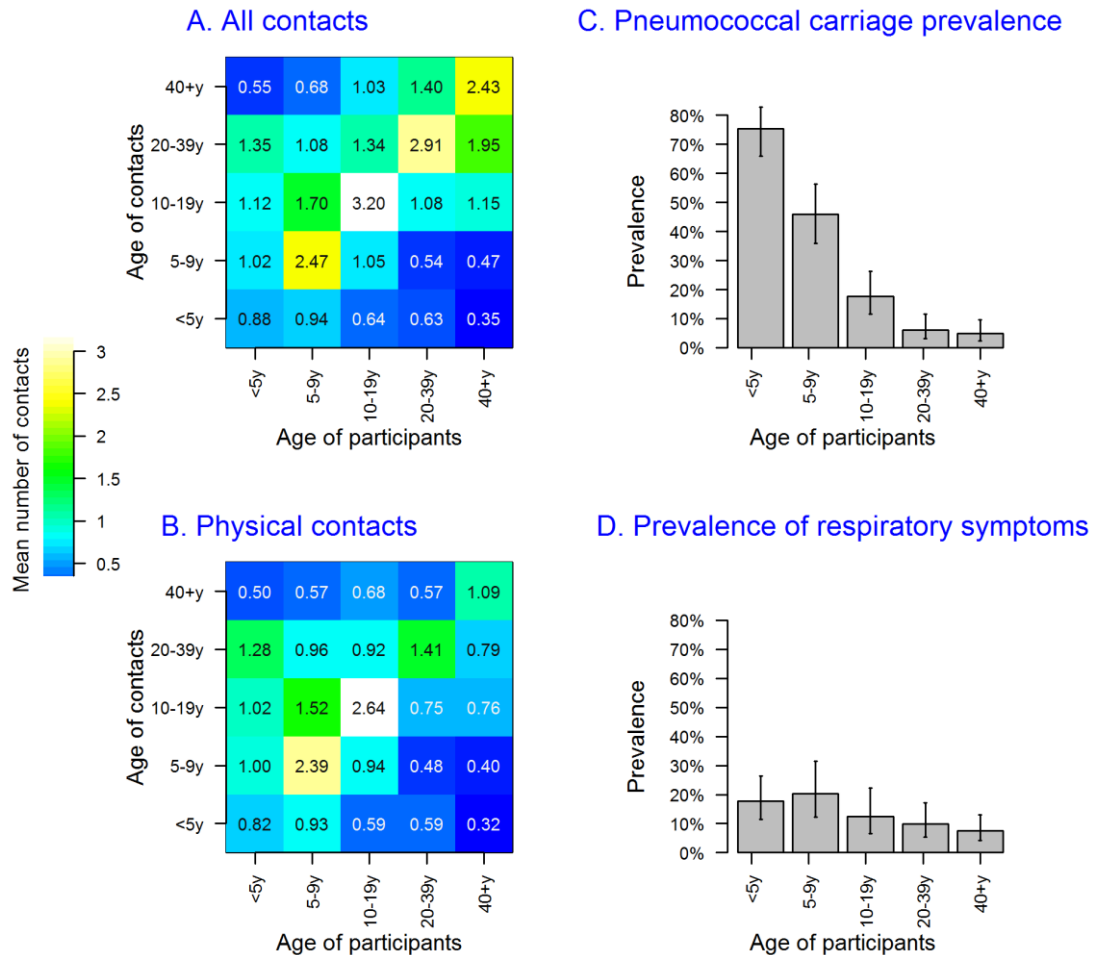
Legend: Contact matrices and prevalence of *S.pneumoniae* carriage (Panel A) and Acute Respiratory Symptoms (Panel B). In panels C and D, the height of each bar corresponds to the point prevalence and the error bar represents the 95% Confidence Interval

The prevalence of pneumococcal carriage was strongly age dependent, decreasing from 75% in

children <5 years to 46% among 5 – 9 year olds, 17% in 10 – 19 year olds, and further

decreasing to 8% and 7% among 20-39 years and ≥40 years old respectively (Figure 16C).

Overall, 72 (13%) people reported having suffered from ARS in the two weeks prior to the

survey. The prevalence of ARS varied much less with age than that of carriage (Figure 16D),

ranging from 20% among 5 to 9 year olds to 8% among ≥40 years old. There was no sex difference in the prevalence of carriage (age-adjusted RR for males 0.97, P=0.814) or ARS (age-adjusted RR for males 1.40, P=0.134).

The Table 7 below provides the crude risk ratios by age for prevalence of carriage, and acute respiratory symptoms (ARS).

Table 7: Crude risk ratios by age for prevalence of carriage, and acute respiratory symptoms (ARS).

Age group	Prevalence	Risk ratio (95%CI)
NP Carriage		
<5 years	75.1%	1.00
5 – 9 years	45.9%	0.61 (0.48 ; 0.77)
10 – 19 years	19.3%	0.24 (0.16 ; 0.37)
20 – 39 years	6.1%	0.08 (0.04 ; 0.15)
≥ 40 years	4.8%	0.06 (0.03 ; 0.13)
ARS		
<5 years	17.4%	1.00
5 – 9 years	20.3%	1.16 (0.63 ; 2.12)
10 – 19 years	10.3%	0.68 (0.32 ; 1.46)
20 – 39 years	9.9%	0.56 (0.28 ; 1.15)
≥ 40 years	7.5%	0.43 (0.22 ; 0.85)

Carriage and ARS were poorly correlated (Pearson's correlation coefficient R=0.09, ranging from -0.08 to 0.29 by age group), and there was no evidence that the risk of ARS was higher among carriers compared to non-carriers (age-adjusted relative risk (RR) 1.06 (95%CI 0.85 – 1.33)).

5.3.2. Social contacts as a risk for pneumococcal carriage or ARS

Overall, the mean number of contacts among carriers was significantly higher than non-carriers, and this observation was consistent across age groups, although most differences were not statistically significant due to small numbers (Figure 17). In particular, carriers had more physical contacts (Figure 17A). This pattern was also consistent for ARS, with the exception of 5 – 9 year olds in whom the mean number of contacts among individuals with ARS was lower for physical contacts (Figure 17B).

In univariable analysis, the risk of carriage increased with all contacts, household contacts, contacts ≥1 hour long and physical contacts. The latter had the largest effect size, with a 13% increased risk for each additional contact reported by participants (Table 8). Physical contacts and contacts ≥1 hour were strongly correlated (R = 0.76), particularly among children <5 years (R=0.85) and children aged 5 – 9 years (R=0.87), hence their effect could not be disentangled, whereas the correlation between physical contacts and household contacts was moderate (R=0.61, ranging from 0.47 to 0.70 between age groups).

After age adjustment, physical contacts or contacts ≥1 hour remained most significantly associated with carriage, with a 6% increased risk (95%CI 2 – 10%) for each unit increase in the number of reported contacts (Table 8). We found good evidence that the number of household contacts increased the risk as well (Table 8). There was no confounding effect by other covariates and models were therefore only adjusted for age (Supporting Information).

After age adjustment, physical contacts or contacts ≥1 hour remained most significantly associated with carriage, with a 6% increased risk (95%CI 2 – 10%) for each unit increase in the number of reported contacts (Table 8). We found good evidence that the number of household contacts increased the risk as well (Table 8). There was no confounding effect by other covariates and models were therefore only adjusted for age (Supporting Information).

Figure 17: Mean number of contacts by age group and nasopharyngeal carriage status (Panel A) or ARS status (Panel B)
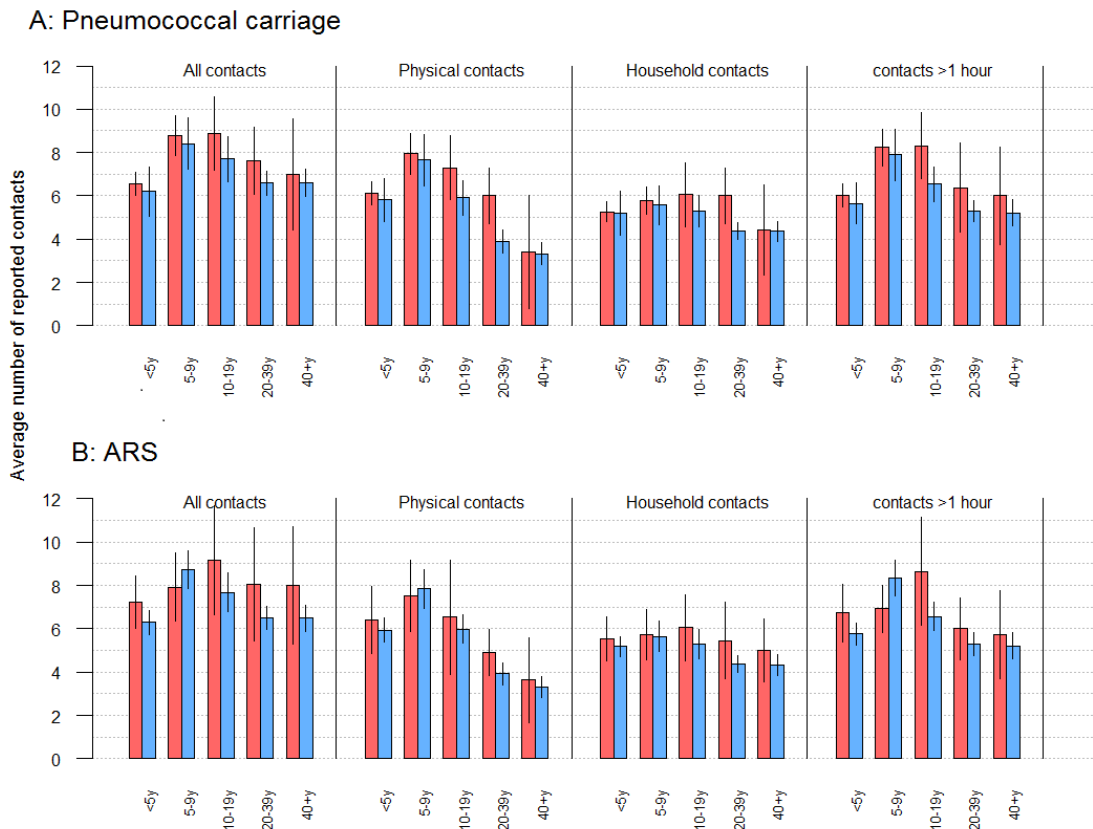
Legend: The graph shows the mean number of contacts among carriers (red, panel A), non-carriers (blue, panel A), individuals reporting ARS (red, panel B), and individuals without ARS (blue, panel B). This is shown by age group and for all contacts, physical contacts, household contacts and contacts lasting over 1 hour.

An increase in physical, household and long (≥ 1h) contacts, were also associated with an increased risk of ARS in univariable analysis. There was little or no evidence of a confounding effect of age, given the more constant prevalence of ARS across age groups (Table 8). Unlike pneumococcal carriage, however, the relative risk was more constant across types of contacts, and household contacts rather than physical contacts were most strongly associated with a risk for ARS, with a risk increase of 9% (1 – 18%) for each additional reported contact.

Table 8: Relative risk of *S.pneumoniae* carriage and ARS by frequency of contact

Contact types	Mean contacts		Crude RR (95%CI)	Age -adjusted RR (95%CI)
NP carriage	**NP+**	**NP-**		
All contacts	8.09	7.53	**1.04 (1.00; 1.07)**	**1.03 (1.00; 1.07)**
Physical contacts	6.83	6.09	**1.13 (1.09; 1.17)**	**1.06 (1.02; 1.09)**
Non-physical contacts	1.25	1.43	0.78 (0.72; 0.86)	0.97 (0.91; 1.05)
Household contacts	5.64	5.31	**1.09 (1.03; 1.14)**	1.04 (0.99 ; 1.10)
Non-household contacts	2.22	2.45	0.98 (0.93; 1.04)	1.03 (0.98; 1.09)
Contacts≥ 1hour	7.50	6.56	**1.08 (1.05; 1.12)**	**1.06 (1.02; 1.10)**
Contacts <1h	2.38	2.62	0.81 (0.73; 0.92)	0.94 (0.85; 1.03)
ARS	**ARS+**	**ARS-**		
All contacts	8.18	7.63	**1.07 (1.02; 1.12)**	**1.07 (1.02; 1.13)**
Physical contacts	6.10	6.26	**1.09 (1.03; 1.15)**	**1.07 (1.00; 1.14)**
Non-physical contacts	2.07	1.36	1.00 (0.89; 1.10)	1.06 (0.95; 1.18)
Household contacts	5.70	5.31	**1.10 (1.02; 1.19)**	**1.09 (1.01; 1.18)**
Non-household contacts	2.47	2.31	1.04 (0.98; 1.11)	1.05 (0.98; 1.13)
Contacts ≥ 1hour	7.33	6.74	**1.07 (1.01; 1.13)**	**1.07 (1.00; 1.14)**
Contacts <1 hour	3.14	2.58	1.05 (0.96; 1.15)	1.08 (0.99; 1.18)

RR in bold: Associations significant at P<0.05. RR= Risk Ratio. CI= Confidence Interval. NP+= pneumococcal carrier, NP - = non-carrier, ARS+ = with ARS, ARS- = without ARS

Next, we analysed whether the number of casual contacts (i.e. contacts lasting < 5 minutes) was associated with either pneumococcal carriage or ARS. We found no evidence that the prevalence of pneumococcal carriage or the risk of ARS were associated with reporting higher levels of casual contacts, as shown in Table 9. Given the small numbers of individuals reporting ≥30 casual contacts, we pooled the 56 individuals reporting ≥20 casual contacts into one category.

In univariable analysis the risk of pneumococcal carriage was higher in the group of 78 participants who did not know how many casual contacts they may have had. This was due to the higher proportion of children <5 years and aged 5 – 9 years in that group, however, after age-adjustment, there was no evidence that the risk of pneumococcal carriage was higher in that group. Similarly, a higher number of casual contacts was associated with increased risk of ARS, and with no confounding effect of age or other variables on the estimates.

We found no confounding effect on casual contacts on the relative risk of carriage or ARS as a function of the frequency of reported 'ordinary' contacts. Finally, we explored the characteristics of age-specific mixing patterns by ARS or carriage status in greater detail, with a focus on physical contacts and household contacts.

Table 9: Relative risk (RR) of pneumococcal carriage and ARS by level of reported number of daily casual contacts

Number	N	Pneumococcal carriage		Acute Respiratory Symptoms (ARS)	
		Crude RR (95% CI)	Age-adjusted (95% CI)	Crude RR (95% CI)	Age-adjusted RR (95% CI)
0 – 9	315	ref	ref	ref	ref
10 – 19	119	0.82 (0.52; 1.27)	0.80 (0.55; 1.75)	1.02 (0.65; 1.49)	1.00 (0.65; 1.54)
≥20	56	0.81 (0.44; 1.49)	1.19 (0.70; 2.04)	0.51 (0.21; 1.27)	0.56 (0.22; 1.39)
Not	76	1.52 (1.00; 2.31)	1.07 (0.77; 1.49)	1.01 (0.45; 2.28)	0.88 (0.38; 2.00)

In bold: Associations significant at P<0.05. RR= Risk Ratio. CI= Confidence Interval

We computed the ratio of the mean number of contacts within and between age groups among carriers compared to non-carriers, and among individuals with ARS compared to those without. The average number of physical encounters within and between age groups tended to be higher for carriers than non-carriers in most instances, albeit with substantial uncertainty owing to small numbers (Figure 18). Carriers reported more contacts with children < 5 years, and particularly adult carriers who reported on average more than twice as many physical contacts with children under five than non-carriers (Figure 18).

Figure 18: Ratio of the mean reported number of *physical* contacts between pneumococcal carriers and non-carries (panel A) and individuals with ARS compared to those without ARS (Panel B).
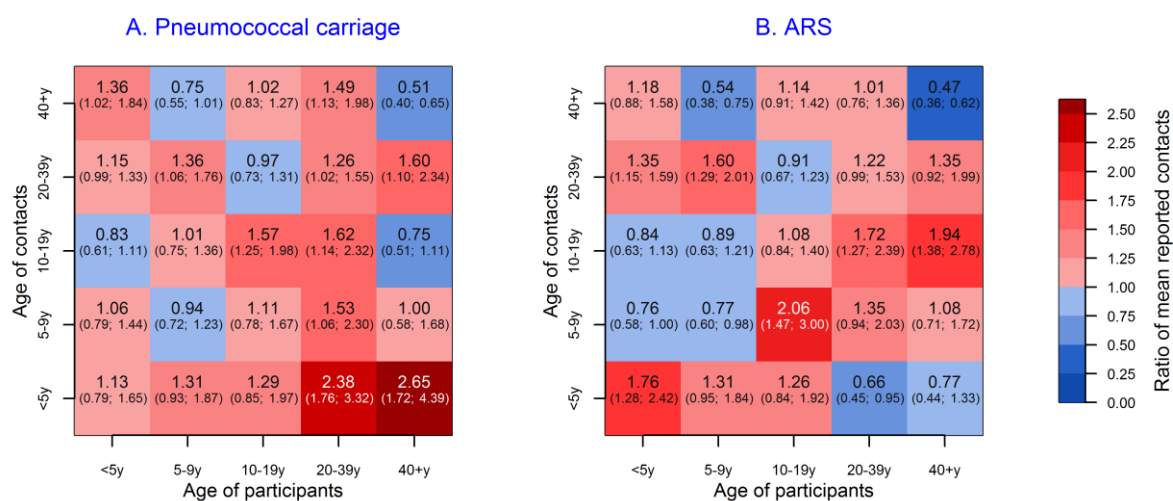
Legend: The numbers represent the point estimate of the ratio and in brackets the 95% confidence bounds. Blue are cells with a point estimate <1 and in red with point estimates >1.

Given that most of such contacts occurred within the household, the effect of physical and household contacts with children <5 years was indistinguishable, whereas for older age groups, the association with physical contact was stronger than that with household contacts (Figure 18).

Similar findings were seen for ARS, however, unlike for pneumococcal carriage, symptomatic adults did not have more contacts with young children than asymptomatic ones. Results based on absolute differences in the mean number of contacts rather than relative means are displayed in Figure S8 (Supporting Information), showing the similar associations than the reported ratios, but providing a quantified difference in mean number of contacts instead.

5.4. Discussion

Our study provided a unique opportunity to explore whether and how social contact patterns are associated with someone's risk of acute respiratory infection in a mostly rural East African setting. Our results show that people who tend to have more frequent close contacts are more

likely to be pneumococcal carriers or to report acute respiratory symptoms, irrespective of their age. In contrast, we found that less intimate or short casual contacts were not associated with someone's infection risk, suggesting that social contacts important for transmission are close interpersonal encounters.

Existing evidence on the association between contact patterns and risk of infection is mostly 'ecological', with very few studies based on individual-level data. Using data of influenza A/H1N1 seroconversion in Hong Kong coupled to social contact data in the same population, Kwok et al. [16] showed that age rather than social contact patterns were the main driver of the individual risk of infection in that setting, and further work by Kucharski et al. [17] on the same data further supported the finding that someone's risk of infection is related to the average mixing pattern within their age group rather than their reported number of contacts. However, the validation of the 'social contact hypothesis' with such data remains difficult, due to challenges in accounting for pre-existing levels of protective immunity in the population, assumptions around stability of behaviour between the capture of social contact data and that of infection (or disease) endpoints, as well as challenges in capturing influenza infection events based on serological data only [16, 17].

Studying nasopharyngeal carriage of *S.pneumoniae* as the main biological endpoint enabled us to address many of these issues, for several reasons. First, the high prevalence of carriage provided the statistical power required to study individually-matched acquisition risk in our study. Second, a person's individual behaviour is unlikely to be affected by carriage given that the vast majority of episodes remain asymptomatic, in contrast to symptomatic respiratory infections during which social behaviour might change as a result of illness [22]. Most carriage episodes in our study were asymptomatic, and our decision to assess the risk for ARS and NP carriage separately stemmed from that observation. In addition, given that duration of carriage

is relatively short – at most 3 months in young children and no longer than a few weeks in adults [26] – , and as no difference in our survey was observed in the mean number of contacts between days of the weeks and between survey weeks, it is reasonable to assume that contact patterns measured on a given day reflect contact patterns around the time of pneumococcal acquisition. Finally, as colonisation results in weak protective immune responses and limited reduction in serotype-specific reacquisition risk [23, 27], and with over 90 circulating serotypes, natural acquired immunity is unlikely to distort the association between carriage and social contact patterns, unlike immunizing infection for which individuals with more frequent risk of infection due to their social contact patterns are also more likely to be immune.

In contrast, self-reported respiratory symptoms may be influenced by factors such as behaviour change in illness [22] or immunity. The definition itself also presents limitations; although respiratory viruses such as the Respiratory Syncytial Virus (RSV), adenovirus, parainfluenza and influenza viruses are likely to account for a large proportion of ARS cases [28], the definition was used as a non-specific proxy for acute respiratory infection and may have captured other infectious and non-infectious conditions. Notwithstanding such caveats, results for ARS were similar to those for nasopharyngeal carriage, but also showed a more consistent association with all types of non-casual contacts, rather than physical or contacts of long duration only. This suggests that the definition may have mostly captured acute infections, and also provides further evidence that the number of close interpersonal contacts, and particularly household contacts, plays a role in the transmission of acute respiratory pathogens.

One of the striking features of our analysis of the relative number of mean contacts between age groups is that adults colonised with *S.pneumoniae* reported more than twice as many close contacts with children under five than non-carriers, and to a lesser extent also more proportionally more contact with adults. The higher number of contacts with children is in agreement with observational and modelling studies showing that pneumococcal carriage risk

increases with household size and with the number of children <5 years in the household [29]. It also supports the finding that carriage acquisition in adults occurs mostly within the household as a result of contact with young children who are drivers of infection [30]. In contrast, however, we found that adults with ARS reported an equal or lower number of contacts with children <5 years than asymptomatic adults. This likely reflects the epidemiological differences between carriage and ARS, given the very high prevalence of carriage among <5 year olds and the much less marked difference in age-specific prevalence of ARS across age groups, in addition to other potential factors such as acquired immunity among adults more frequently exposed to very young children. Some of the specific details of the contact patterns and differences between *S.pneumoniae* and ARS are harder to interpret, due in part to wide statistical uncertainty, particularly among adults in whom the number of carriers or symptomatic individuals is small. However, overall the results from our analysis support the general finding that increased close contacts are associated with higher risk of ARS or pneumococcal carriage.

There is still much debate about how respiratory pathogens are transmitted from person to person; whether through close direct or indirect physical contact, through large droplet transmission at close range (<1 meter), or through aerosolized particles floating over longer distances, particularly in poorly ventilated indoor settings [31, 32]. It is likely that many pathogens can be transmitted through a combination of these routes. Yet the contribution of each mechanism remains uncertain [31, 32]. It is generally assumed that the main transmission route of *S.pneumoniae* is through direct contact [33], as well as indirectly through shared glasses or bottles [34]. Analogously, for other colonising bacteria such as *N.meningitidis*, close contact and intimate kissing are known risk factors among teenagers and young adults [35]. Similarly, it is believed that influenza and other respiratory viruses are primarily transmitted

through direct contact or contact with large droplet transmission at close range rather than aerosolized particles [31].

Although our objective was not to demonstrate transmission, our findings strongly support that direct close interpersonal contact is an important mode of transmission for pathogenic bacteria of the nasopharynx as well as respiratory viruses, and strengthens the scientific evidence for public health measures based upon these assumptions, such as hand washing campaigns or chemoprophylaxis of close contacts of cases of meningococcal meningitis to name a few. The strong association of ARS with household and other close contacts, independently from physical touch, might suggest that other mechanisms such as indirect transmission through fomites or aerosol transmission may play a role. Elucidating the contribution of such factors in this context would be an important question for future research.

Our findings have also several implications for infectious disease research. Contact structures are central to transmission models, and appropriate assumptions about what type of contact drives infectious disease transmission are essential. Our results suggest that the parameterisation of transmission models of *S.pneumoniae* and similar pathogens using mixing matrices based on physical or another measure of *close* interpersonal contact would more likely capture relevant contact patterns than those based on any type of social encounter. This has also implications for the design of contact studies, particularly in low-income settings given the scarcity of published data currently available and the need to collect additional data from many more settings [7, 8]. In contrast to diary-based approaches that have been used by many [2, 4, 7, 8], the study design here was relatively simple and only involved a single face-to-face interview. A drawback of such retrospective approach is the lack of detailed information about very short (i.e. 'casual') contacts, as such information was deemed unreliable – and this is further supported by evidence that short contacts tend to be inconsistently recorded even in

prospective diary-based approached [1]. However, assuming that short contacts do not account for much of the transmission, as our study suggests, a more simple retrospective design is a potential attractive option in other settings where contact data are lacking, and in which data collection through more comprehensive diary-based approaches may be difficult to implement.

There are some additional limitations to our study. We were unable to explore other potential confounding factors, such as bed share, ventilation, indoor smoke or hand washing [36], and the contribution of such factors should be explored further. Among other important factors, it is worth mentioning that our results were not impacted by pneumococcal vaccination, as the vaccine had not been introduced in the study area at the time. Moreover, it remains unclear to what extent our results can be generalised to any acute respiratory infection. For example, factors such as ventilation and airflow may be of particular importance for aerosolized transmission of pathogens such as mycobacteria [7], compared to influenza or *S.pneumoniae* [31, 32], and whether household or physical contacts reflect contact patterns important for aerosolized transmission remains uncertain [7]. Finally, while we found a strong association at the individual level, our study does not demonstrate causality.

However, in the absence of detailed longitudinal data on acquisition events between every single contact —which would be challenging and possibly unrealistic to obtain— our findings provide consistent evidence of a 'dose-response' association at the individual level between close social encounters and acquisition risk for respiratory pathogens, and therefore provides robust support both for the social contact hypothesis, and for research and policy work based upon this hypothesis.

Acknowledgments

Conflict of interest

The authors declare they have no conflict of interest.

Funding

References

1.   Smieszek T, Burri EU, Scherzinger R, Scholz RW. Collecting close-contact social mixing data with contact diaries: reporting errors and biases. Epidemiology and infection. 2012;140(4):744-52. doi: 10.1017/S0950268811001130. PubMed PMID: 21733249.

2.   Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine. 2008;5(3):e74. doi: 10.1371/journal.pmed.0050074. PubMed PMID: 18366252; PubMed Central PMCID: PMC2270306.

3.   Eubank S, Guclu H, Kumar VS, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. Nature. 2004;429(6988):180-4. doi: 10.1038/nature02541. PubMed PMID: 15141212.

4.   Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. American journal of epidemiology. 2006;164(10):936-44. doi: 10.1093/aje/kwj317. PubMed PMID: 16968863.

5.   Horby P, Pham QT, Hens N, Nguyen TT, Le QM, Dang DT, et al. Social contact patterns in Vietnam and implications for the control of infectious diseases. PloS one. 2011;6(2):e16965. doi: 10.1371/journal.pone.0016965. PubMed PMID: 21347264; PubMed Central PMCID: PMC3038933.

6.   Grijalva CG, Goeyvaerts N, Verastegui H, Edwards KM, Gil AI, Lanata CF, et al. A household-based study of contact networks relevant for the spread of infectious diseases in the highlands of Peru. PloS one. 2015;10(3):e0118457. doi: 10.1371/journal.pone.0118457. PubMed PMID: 25734772; PubMed Central PMCID: PMC4348542.

7.   Johnstone-Robertson SP, Mark D, Morrow C, Middelkoop K, Chiswell M, Aquino LD, et al. Social mixing patterns within a South African township community: implications for respiratory disease transmission and control. American journal of epidemiology. 2011;174(11):1246-55. doi: 10.1093/aje/kwr251. PubMed PMID: 22071585; PubMed Central PMCID: PMC3224253.

8.   Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. PloS one. 2014;9(8):e104786. doi: 10.1371/journal.pone.0104786. PubMed PMID: 25127257; PubMed Central PMCID: PMC4134222.

9.   Fu YC, Wang DW, Chuang JH. Representative contact diaries for modeling the spread of infectious diseases in Taiwan. PloS one. 2012;7(10):e45113. doi: 10.1371/journal.pone.0045113. PubMed PMID: 23056193; PubMed Central PMCID: PMC3463600.

10.   Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and urban areas of southern China. Proceedings Biological sciences / The Royal Society. 2014;281(1785):20140268. doi: 10.1098/rspb.2014.0268. PubMed PMID: 24789897; PubMed Central PMCID: PMC4024290.

11.   Edmunds WJ, O'Callaghan CJ, Nokes DJ. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proceedings Biological sciences / The Royal Society. 1997;264(1384):949-57. doi: 10.1098/rspb.1997.0131. PubMed PMID: 9263464; PubMed Central PMCID: PMC1688546.

12.   Halloran ME, Ferguson NM, Eubank S, Longini IM, Jr., Cummings DA, Lewis B, et al. Modeling targeted layered containment of an influenza pandemic in the United States. Proceedings of the National Academy of Sciences of the United States of America.

2008;105(12):4639-44. doi: 10.1073/pnas.0706849105. PubMed PMID: 18332436; PubMed Central PMCID: PMC2290797.

13.   Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modelling study. PLoS medicine. 2013;10(10):e1001527. doi: 10.1371/journal.pmed.1001527. PubMed PMID: 24115913; PubMed Central PMCID: PMC3793005.

14.   Santermans E, Goeyvaerts N, Melegaro A, Edmunds WJ, Faes C, Aerts M, et al. The social contact hypothesis under the assumption of endemic equilibrium: Elucidating the transmission potential of VZV in Europe. Epidemics. 2015;11:14-23. doi: 10.1016/j.epidem.2014.12.005. PubMed PMID: 25979278.

15.   Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ. What types of contacts are important for the spread of infections?: using contact survey data to explore European mixing patterns. Epidemics. 2011;3(3-4):143-51. doi: 10.1016/j.epidem.2011.04.001. PubMed PMID: 22094337.

16.   Kwok KO, Cowling BJ, Wei VW, Wu KM, Read JM, Lessler J, et al. Social contacts and the locations in which they occur as risk factors for influenza infection. Proceedings Biological sciences / The Royal Society. 2014;281(1789):20140709. doi: 10.1098/rspb.2014.0709. PubMed PMID: 25009062; PubMed Central PMCID: PMC4100506.

17.   Kucharski AJ, Kwok KO, Wei VW, Cowling BJ, Read JM, Lessler J, et al. The contribution of social behaviour to the transmission of influenza A in a human population. PLoS pathogens. 2014;10(6):e1004206. doi: 10.1371/journal.ppat.1004206. PubMed PMID: 24968312; PubMed Central PMCID: PMC4072802.

18.   Rohani P, Zhong X, King AA. Contact network structure explains the changing epidemiology of pertussis. Science. 2010;330(6006):982-5. doi: 10.1126/science.1194134. PubMed PMID: 21071671.

19.   O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet. 2009;374(9693):893-902. doi: 10.1016/S0140-6736(09)61204-6. PubMed PMID: 19748398.

20.   Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375(9730):1969-87. doi: 10.1016/S0140-6736(10)60549-1. PubMed PMID: 20466419.

21.   Cohen AL, Hyde TB, Verani J, Watkins M. Integrating pneumonia prevention and treatment interventions with immunization services in resource-poor countries. Bulletin of the

World Health Organization. 2012;90(4):289-94. doi: 10.2471/BLT.11.094029. PubMed PMID: 22511825; PubMed Central PMCID: PMC3324866.

22. Van Kerckhove K, Hens N, Edmunds WJ, Eames KT. The impact of illness on social networks: implications for transmission and control of influenza. American journal of epidemiology. 2013;178(11):1655-62. doi: 10.1093/aje/kwt196. PubMed PMID: 24100954; PubMed Central PMCID: PMC3842903.

23. Prevaes SM, van Wamel WJ, de Vogel CP, Veenhoven RH, van Gils EJ, van Belkum A, et al. Nasopharyngeal colonization elicits antibody responses to staphylococcal and pneumococcal proteins that are not associated with a reduced risk of subsequent carriage. Infection and immunity. 2012;80(6):2186-93. doi: 10.1128/IAI.00037-12. PubMed PMID: 22451514; PubMed Central PMCID: PMC3370583.

24. Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, et al. Standard method for detecting upper respiratory carriage of Streptococcus pneumoniae: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. Vaccine. 2013;32(1):165-79. doi: 10.1016/j.vaccine.2013.08.062. PubMed PMID: 24331112.

25. Rogers WH. Regression standard errors in clustered samples. Stata Technical Bulletin 1993;13:19–23.

26. Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JA. The descriptive epidemiology of Streptococcus pneumoniae and Haemophilus influenzae nasopharyngeal carriage in children and adults in Kilifi district, Kenya. The Pediatric infectious disease journal. 2008;27(1):59-64. doi: 10.1097/INF.0b013e31814da70c. PubMed PMID: 18162940; PubMed Central PMCID: PMC2382474.

27. Lipsitch M, Abdullahi O, D'Amour A, Xie W, Weinberger DM, Tchetgen Tchetgen E, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. Epidemiology. 2012;23(4):510-9. doi: 10.1097/EDE.0b013e31824f2f32. PubMed PMID: 22441543; PubMed Central PMCID: PMC3670084.

28. Berkley JA, Munywoki P, Ngama M, Kazungu S, Abwao J, Bett A, et al. Viral etiology of severe pneumonia among Kenyan infants and children. Jama. 2010;303(20):2051-7. doi: 10.1001/jama.2010.675. PubMed PMID: 20501927; PubMed Central PMCID: PMC2968755.

29. Farida H, Severin JA, Gasem MH, Keuter M, Wahyono H, van den Broek P, et al. Nasopharyngeal carriage of Streptococcus pneumonia in pneumonia-prone age groups in Semarang, Java Island, Indonesia. PloS one. 2014;9(1):e87431. doi:

10.1371/journal.pone.0087431. PubMed PMID: 24498104; PubMed Central PMCID: PMC3909120.

30.   Melegaro A, Gay NJ, Medley GF. Estimating the transmission parameters of pneumococcal carriage in households. Epidemiology and infection. 2004;132(3):433-41. PubMed PMID: 15188713; PubMed Central PMCID: PMC2870123.

31.   Brankston G, Gitterman L, Hirji Z, Lemieux C, Gardam M. Transmission of influenza A in human beings. The Lancet Infectious diseases. 2007;7(4):257-65. doi: 10.1016/S1473-3099(07)70029-4. PubMed PMID: 17376383.

32.   In: Larson EL, Liverman CT, editors. Preventing Transmission of Pandemic Influenza and Other Viral Respiratory Diseases: Personal Protective Equipment for Healthcare Personnel: Update 2010. Washington (DC)2011.

33.   Ferreira DM, Neill DR, Bangert M, Gritzfeld JF, Green N, Wright AK, et al. Controlled human infection and rechallenge with Streptococcus pneumoniae reveals the protective efficacy of carriage in healthy adults. American journal of respiratory and critical care medicine. 2013;187(8):855-64. doi: 10.1164/rccm.201212-2277OC. PubMed PMID: 23370916; PubMed Central PMCID: PMC3707375.

34.   Levine H, Balicer RD, Zarka S, Sela T, Rozhavski V, Cohen D, et al. Dynamics of pneumococcal acquisition and carriage in young adults during training in confined settings in Israel. PloS one. 2012;7(10):e46491. doi: 10.1371/journal.pone.0046491. PubMed PMID: 23056322; PubMed Central PMCID: PMC3466294.

35.   MacLennan J, Kafatos G, Neal K, Andrews N, Cameron JC, Roberts R, et al. Social behavior and meningococcal carriage in British teenagers. Emerging infectious diseases. 2006;12(6):950-7. PubMed PMID: 16707051; PubMed Central PMCID: PMC3373034.

36.   Tumwesigire SG, Barton T. Environmental risk factors for acute respiratory infections among children of military personnel in Uganda. East African medical journal. 1995;72(5):290-4. PubMed PMID: 7555884.

5.5. Supporting Information

5.5.1. Additional Figures

Figure S8: Difference in the mean number of *physical* contact between pneumococcal carriers and non-carriers (Panel A) and between individuals suffering and not suffering from respiratory symptoms (Panel B)
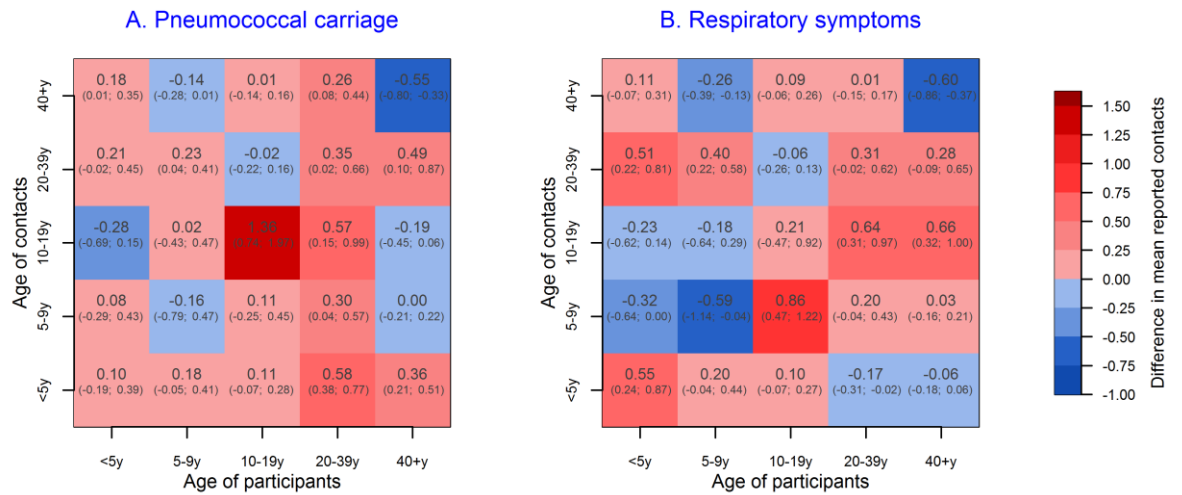
Legend: Matrices of the mean difference between the mean number of contacts among pneumococcal carriers and non-carriers (panel A), and individuals with ARS compared to non-ARS (panel B). The numbers represent the point estimate and the lower and upper bound of the 95% credible interval are shown inside the brackets.

5.5.2. Analysis and model comparison

5.5.2.1. *Univariable analysis*

In univariable analysis we explored the risk of pneumococcal carriage or ARS as a function of the following covariates: age, sex, and all types of ordinary contacts (all, physical (skin-to-skin), household, long (≥1h), non-physical, non-household, short (<1h)). We used a log-binomial regression model through a generalized linear model with a robust variance estimator, and inclusion probability weights by age group to capture the differences in inclusion by age at the sampling stage.

We considered contacts as linear predictors, but tested departure from linearity based on BIC values and likelihood ratio tests comparing models with contacts as linear variables and as categorical variables. On that basis there was no evidence of departure from linearity (BIC larger in all cases and P-value of likelihood ratio tests >P=0.10)

5.5.3. Multivariable analysis

5.5.3.1.   Pneumococcal carriage

Age was first added to each model as an a priori confounding variable. Age-adjustment improved all models of pneumococcal risk as a function of all other covariates, leading to substantial changes in measures of effect (based on a lower BIC and a likelihood ratio test P-value <0.05).

We modelled the association between age-adjusted covariates and carriage through the following models: $\log(P_{AX}) = \beta_0 + \beta_1 Age + \beta_2 X$, where A is Age and X was the covariate of interest (sex, or any type of ordinary contact (all, physical, etc.))

Only all contacts, physical contacts, household contacts and long (≥ 1h) contacts were associated with carriage after age adjustment. There was a strong correlation between physical contacts and contacts lasting more than one hour (Pearon's correlation coefficient R=0.76), particularly in children <10 years of age (R>0.85), hence those covariates were not added together in a model due to collinearity. For household contacts correlation was moderate with physical contacts (R=0.61) or long contacts (R=0.64) and we therefore explored models including both physical contacts and household contacts, as well as contacts ≥1h and household contacts, in order to explore their potential confounding effect.

Table S4 provides the BIC values for the different models tested for the association with physical, household and long contacts. We found that model included a single type of contact as covariate performed better than those including two or more, hence only age-adjusted estimates were presented in the main paper.

Table S4: Comparing the BIC for models of pneumococcal carriage risk as a function of physical, household and long contact.

Model [covariates]	BIC
$\log(P_P) = \beta_0 + \beta_1 P$ [P=Physical contacts]	-3218
$\log(P_H) = \beta_0 + \beta_1 H$ [H=Household contacts]	-3199
$\log(P_L) = \beta_0 + \beta_1 L$ [L=Long (≥1h) contacts]	-3202
$\log(P_{AP}) = \beta_0 + \beta_1 A + \beta_2 P$ [A=Age, Physical contacts]	-3312
$\log(P_{AP}) = \beta_0 + \beta_1 A + \beta_2 H$ [Age, Household contacts]	-3310
$\log(P_{AP}) = \beta_0 + \beta_1 A + \beta_2 L$ [Age, Long (≥1h) contacts]	-3312
$\log(P_{APH}) = \beta_0 + \beta_1 A + \beta_2 P + \beta_3 H$ [Age, Physical, Long]	-3306
$\log(P_{ALH}) = \beta_0 + \beta_1 A + \beta_2 L + \beta_3 H$ [Age, Long, Household]	-3306

In bold: best models

5.5.3.2.   ARS

The same approach as for pneumococcal carriage was taken. Again, age was considered as the main a priori confounding factor, given the differences in both carriage frequency and ARS prevalence by age. However, there was no improvement in model performance (no change or increase in BIC and non-significant likelihood ratio test) and little or no change in point estimates. Adding other variables to the model resulted in worse model performance.

Table S5 shows the different models explored for covariates associated with ARS at P<0.05 in univariable analysis (Physical contacts, household contacts, and contacts >1h).

Table S5: Models explored for ARS

Model [covariates]	BIC
$\log(P_P) = \beta_0 + \beta_1\,P$ [P=Physical contacts]	-3288
$\log(P_H) = \beta_0 + \beta_1\,H$ [H=Household contacts]	-3286
$\log(P_L) = \beta_0 + \beta_1\,L$ [L=Long (≥1h) contacts]	-3286
$\log(P_{AP}) = \beta_0 + \beta_1\,A + \beta_2\,P$ [A=Age, Physical contacts]	-3285
$\log(P_{AP}) = \beta_0 + \beta_1\,A + \beta_2\,H$ [Age, Household contacts]	-3286
$\log(P_{AP}) = \beta_0 + \beta_1\,A + \beta_2\,L$ [Age, Long (≥1h) contacts]	-3285
$\log(P_{APH}) = \beta_0 + \beta_1\,A + \beta_2\,P + \beta_3\,H$ [Age, Physical, Long]	-3279
$\log(P_{ALH}) = \beta_0 + \beta_1\,A + \beta_2\,L + \beta_3\,H$ [Age, Long, Household]	-3280

In bold: best models

6. Research paper 5: Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam

Olivier le Polain de Waroux[1] ; W. John Edmunds [1] ; Kensuke Takahashi[2] ; Koya Ariyoshi[2] ; E. Kim Mulholland[1,5]; David Goldblatt[6]; Yoon Hong Choi[3,4] ; Duc-Anh Dang [7] ; Lay Myint Yoshida[2];

Stefan Flasche[1]

1. Department of Infectious Disease Epidemiology department London School of Hygiene and Tropical Medicine, London, United Kingdom

2. Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan

3. Immunisation, Hepatitis and Blood Safety Department, Public Health England, London, United Kingdom

4. Modelling and Economics Unit, Public Health England, London, United Kingdom

5. Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Australia.

6. Institute of Child Health, University College London, London, United Kingdom

7. National Institute of Hygiene and Epidemiology, Hanoi, Vietnam

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Olivier le Polain de Waroux
Principal Supervisor	W John Edmunds
Thesis Title	Epidemiology and transmission dynamics of Streptococcus pneumoniae in low and lower-middle income settings: implications for vaccination strategies

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	NA		
When was the work published?	NA		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

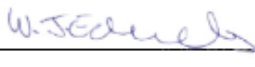|	
Where is the work intended to be published?	BMC Infectious Diseases
Please list the paper's authors in the intended authorship order:	Olivier le Polain de Waroux ; W. John Edmunds ; Kensuke Takahashi ,; Koya Ariyoshi ,; E. Kim Mulholland; David Goldblatt; Yoon Hong Choi; Duc-Anh Dang; Lay Myint Yoshida,; Stefan Flasche
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The candidate defined the research question and developed the model in collaboration with SF and WJE; the candidate parameterised and fitted the model with input from WJE and SF performed the analyses and

	drafted the manuscript. Other co-authors were involved in data collection in the field, data interpretation and provided guidance on study objectives and model parameters. All authors revised the manuscript and approved its final version, now ready for submission

Student Signature: _____   **Date:** 18 November 2016

Supervisor Signature: _____   **Date:** 18 November 2016

Abstract

Introduction

Catch-up campaigns (CCs) at the introduction of the pneumococcal conjugate vaccines (PCVs) may accelerate the impact of PCVs. However, limited vaccine supplies may delay vaccine introduction if additional doses are needed for such campaigns. We studied the relative impact of introducing PCV13 with and without catch-up campaign, and the implications of potential introduction delays

Methods

We used a dynamic transmission applied to the population of Nha Trang, Vietnam. Four strategies were considered: routine vaccination (RV) only, and RV alongside catch-up campaigns (CCs) among <1y olds (CC1), <2y olds (CC2) and <5y olds (CC5). The model was parameterised with data on contact rates from Nha Trang, and was fitted to local carriage data. Post-PCV predictions were based on best estimates of parameters governing post-PCV dynamics, including serotype competition, vaccine efficacy and duration of protection.

Results

Our model predicts elimination of vaccine-type (VT) carriage across all age groups within 10 years of introduction in all scenarios with near-complete replacement by non-VT. Most of the benefit of CCs is predicted to occur within the first 3 years after introduction, with the highest impact at one year, when IPD incidence is predicted to be 11% (95%CrI 9 – 14%) lower than RV with CC1, 25% (21 – 30 %) lower with CC2 and 38% (32 – 46%) lower with CC5.

However, CCs would only prevent more cases of IPD insofar as such campaigns do not delay introduction by more than 31 (95%CrI 30 – 32) weeks with CC1, 58 (53 – 63) weeks with CC2 and 89 (78 – 101) weeks for CC5

Conclusion

CCs are predicted to offer a substantial additional reduction in pneumococcal disease burden over RV alone, if their implementation does not result in much introduction delay. Those findings are important to help guide vaccine introduction in countries that have not yet introduced PCV, particularly in Asia.

6.1. Introduction

Disease due to *Streptococcus pneumoniae* (the pneumococcus) is a leading cause of morbidity and mortality worldwide, disproportionally so in resource-poor settings [1-3].

Ten- and 13-valent pneumococcal conjugate vaccines (PCV10 and PCV13), which cover 10 and 13 of 94 known serotypes [4], are steadily being introduced into the routine immunization programmes of many low and lower-middle income countries with the support from Gavi, the Vaccine Alliance. In 2015, 50 out of 73 countries eligible for Gavi support have introduced the vaccine, and eight additional countries have been approved for introduction and are expected to introduce PCV within the next two years [5]. In 2012 Pakistan was the first Asian country to introduce PCV for routine use, followed by Nepal, Cambodia and Lao PDR [5]. However, PCV has not been introduced into the routine childhood immunisation programme in most of Asia, including Vietnam, and globally, the vaccination coverage of PCV remains low, at 25% [6].

The WHO recommends introducing PCV into childhood immunisation programmes alongside a catch-up campaign (CC) among older children [7]. The aim of such campaign is to provide direct protection to age groups at particular risk of pneumococcal disease, as well as to accelerate the population impact of the vaccine in a timelier fashion through enhanced herd protection [8].

However, the magnitude of the additional impact of different CCs over RV remains unclear. Moreover, CCs have so far not been conducted in most Gavi-supported countries over concerns that they would lead to vaccine introduction delays given limited vaccine supplies [5].

We explored the differential impact on carriage and invasive disease of catch-up campaigns targeting various age groups, through a dynamic compartmental model of disease transmission, and we explored the possible impact of vaccine introduction delay.

6.2.   Methods

The model was applied to the population of Nha Trang (~360,000 inhabitants), an urban and semi-rural area in south-central Vietnam.

6.2.1.   Data

6.2.1.1.   *Nasopharyngeal carriage*

Two surveys were conducted among children aged 0 – 59 months randomly drawn from the population census, with 350 children included in January 2008 [9] and another 350 children in July 2008 [10].

Samples were processed and cultured as per WHO recommendations [11]. Serotyping was done using polymerase chain reaction with 29 specific primer pairs, and did not differentiate between 6A and 6B serotypes [9]. Given that both antigens are included in PCV13 but that PCV10 does not include 6A, we decided to implement our model for PCV13.

The carriage prevalence (and its uncertainty) of VT and NVT among 5-17 year olds and adults (≥18 years) was estimated based on the prevalence and serotype distribution in children <5 years of age, using a random effect meta-regression model [12].

6.2.1.2. Social mixing patterns

We derived the age-specific contact patterns from a survey conducted in the same area in 2010. Study participants were interviewed about their contact on a given day, similarly to the approach undertaken in Uganda as described in Research Paper 3 of the thesis. Participants were first asked to fill in some background information, and were then asked to remember their social encounter occurring in the following 24 hours, before being interviewed about those shortly after. As for other social contact studies, the questionnaire recorded information on the location and type of contacts, including whether or not contacts included physical (i.e. skin-to-skin touch).

In the study, household were randomly selected from all non-touristic urban and semi-rural areas of Nha Trang (i.e. matching the areas where the nasopharyngeal carriage was undertaken)., and one person per household was included in the study. The sample size was based on POLYMOD calculations [13), but was higher to account for urban/rural differences.

The corresponding mixing matrix used in the model was derived as in Melegaro et al. [14]. We parameterised our model based on physical (skin-to-skin) contacts only, given that *S.pneumoniae* is generally assumed to be transmitted through close interpersonal contact [15], in agreement with Research paper 4 of this thesis.

Ethics approval for both the contact and nasopharyngeal surveys was granted by the institutional review board of the London School of Hygiene and Tropical Medicine and the Ethics Commission of National Institute of Hygiene and Epidemiology, Hanoi and the Nagasaki University. A total of 2002 individuals were included in the study.

The mean number of contacts per age group is shown in Figure 19 here below for physical contacts. Further results can be obtained on request.

Figure 19: Mean number for physical contacts between age groups of study participants
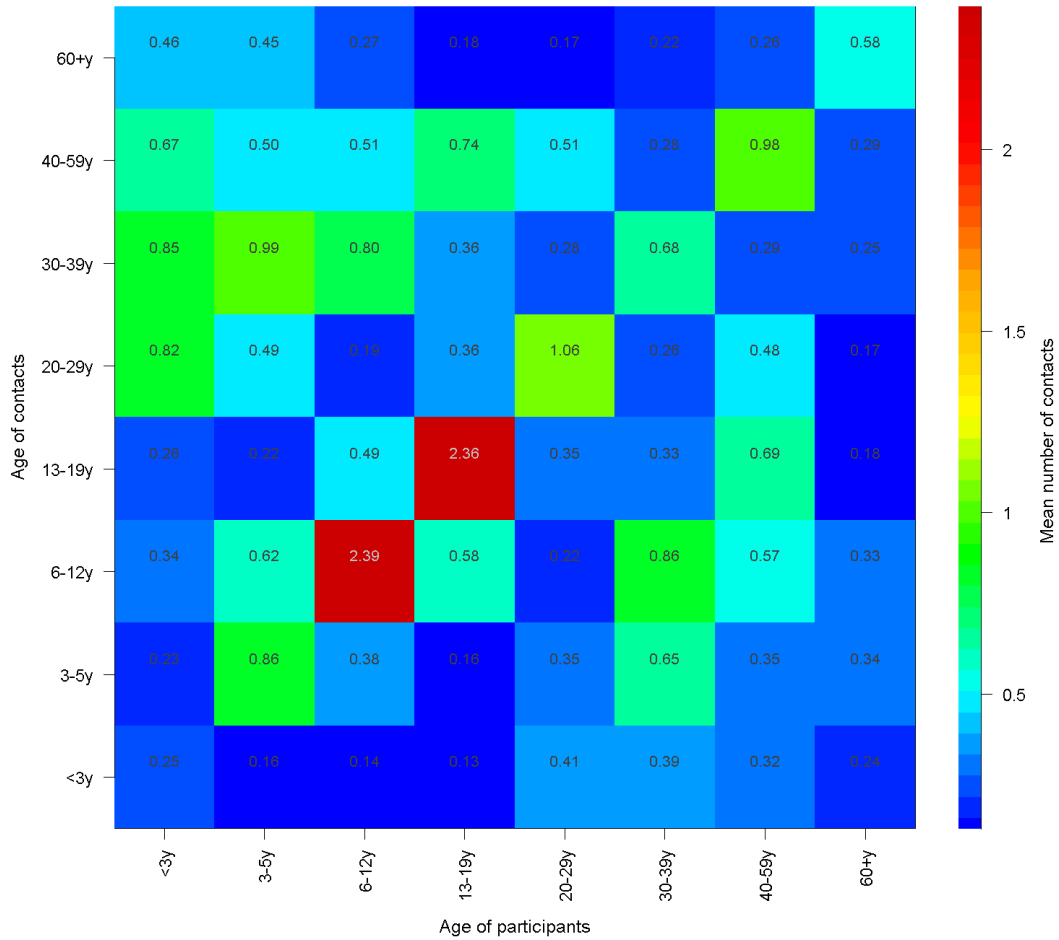
6.2.2. Model structure

We built an age-structured deterministic Susceptible-Infected-Susceptible transmission model of carriage acquisition and clearance, in which we modelled VT jointly and separately from NVT, but allowed for co-colonisation, as in previous models [16]. Details about the model structure and model equations can be found in the Supporting Information (Part 1).

The model comprised of three levels of vaccine-induced immunity; (1) no protection, (2) partial protection and (3) full protection. The latter refers to the efficacy and duration of protection conferred after completion of the infant schedule (i.e. 2 infant doses and a booster at 12 months ('2+1' schedule)) or the completion of a catch-up programme in older children (2 doses in <18 months and 1 dose in ≥18 months). Partial protection was gained form two primary infant doses, or after the first catch-up dose in children aged 12 - 17 months. The difference between full and partial protection relied in the magnitude of vaccine efficacy against carriage (VE_C) and in the duration of protection.

We applied the model to a population of 81 annual age cohorts (0 to 80 years) divided into 52 weekly age bands of 100 individuals. In the calculation of the force of infection the population figure was adjusted to represent the actual population, based on census data.

6.2.3. Model fitting

We fitted the model to pre-vaccination nasopharyngeal carriage data using a Markov Chain Monte-Carlo (MCMC) algorithm, and estimating the age-specific probability of effective transmission in the absence of PCV. For each unique posterior sample we simulated up to 15 years after vaccine introduction.

The model was fitted the model to pre-vaccination nasopharyngeal carriage data using an MCMC algorithm. Estimates of the age-specific probability of effective transmission, which were obtained from the contact and carriage data were updated in the MCMC. Fitting was performed using a Poisson log likelihood, as follows:

$$\log L = \sum_{a=1,n}^{n} (log(vt_a) \times VT_a - vt_a) + \sum_{a=1,n}^{n} (log(nvt_a) \times NVT_a - nvt_a) + constant$$

, where vt_a is the modelled number of VT carriers in age group a, VT_a is the observed number

of VT carriers in age group a, nvt_a is the modelled number of NVT carriers in age group a,

NVT_a is the observed number of NVT in age group a. The constant here is independent of the

modelled number of cases.

The MCMC was run with 100,000 iterations of the chain and for each accepted set of

parameters the post-vaccination simulation up to 15 years after vaccine introduction was

computed, after a burn-in of 20,000 iterations.

Convergence of the log likelihood was first assessed visually and then with the Geweke test of

convergence [49].

In the calculation of the log-likelihood for carriage, model estimates of VT carriage prevalence

included VT carriers and VT-NVT co-colonization given the higher likelihood of the PCR assay

to detect VT than NVT colonies in case of multiple colonization [9].

6.2.4. Model parameters and input

Table 9 displays the value assigned to the different parameters governing transmission and

vaccination. Uncertainty around parameters was taken into account by sampling from their

posterior distribution in the MCMC process.

We obtained the vaccine efficacy against carriage conferring full protection (VE_C^F) and its

uncertainty from a meta-regression model [17]. We estimated the partial efficacy against

carriage (VE_C^P) as 0.78 (95% CI 0.64 – 0.92) that of VE_C^F through a meta-analysis of the relative

risk of VT carriage in 2+0 schedules compared to 3+0 or 2+1 schedules, which included four

trials and eight entry points [18-21]. Further details are provided in the Supporting Information

Part Two.

We assumed an exponential decay function for the waning of vaccine efficacy, as in previous models [16, 22]. We fixed the average duration of protection to 6 years, which best matched the output of an asymptomatic meta-regression model of waning efficacy [17], and analysed the impact of shorter (3 years) and longer (20 years) average protection in sensitivity analyses, based on the 95% credible intervals (CrI) of a model of waning efficacy [17]

Table 10: Parameters used in the model

Parameter	Value	Source*
Competition for carriage acquisition $(C_N$ and $C_V)^\$$	0.1 (SD 0.01)	[16, 32]
Duration of carriage (δ_i)		
0 – 11 months	47.1 days	[49]
12 – 23 months	39.4 days	[49]
24 – 35 months	31.6 days	[49]
36 – 47 months	21.5 days	[49]
48 – 59 months	21.3 days	[49]
5 – 17 years	17.0 days	[22, 50]
18 years and over	18.0 days	[22, 50]
Vaccine efficacy against carriage		
Full protection (VE_{Cf})	63% (95%CrI 53 –73 %)	[17]
Partial protection (VE_{Cp})	0.78 (95%CI 0.64 – 0.92) * VE_{Cf}	Meta-analysis of [18-21] (Supporting Information Part Two)
Mean duration of protection against carriage		
In fully protected (D_F)	6.0 years	[17](Supporting Information Part Two)
In partially protected (D_P)	0.78 (95%CI 0.64 – 0.92) * D_F	Based on same assumptions as for Vaccine efficacy
Vaccine efficacy against IPD (VE_{IPD})	0.80 (0.61 – 0.90)	[51]
Vaccine efficacy against invasiveness (VE_{INV})	0.48 (0.34 – 0.65)	Derived from estimates of [51] and [17] (Supporting Information Part Two)
Mixing matrix	Social contact study (Nha trang)	Calculated as in Melegaro et al [14]

$^\$$Where 0.1 means a force of infection which is 10% that of a situation with no competition

The average duration of protection following particular vaccination was assumed to be 0.78 (95%CI 0.64 – 0.92) that of a complete schedule, as for vaccine efficacy. Details are provided in the Supporting Information (Part Two).

In our base case model we set the vaccination coverage of both routine and catch-up strategies at 90%, in accordance with coverage data from Vietnam for routine immunization programmes and supplementary immunization activities [23].

Carriage estimates were translated into estimates of invasive pneumococcal disease based on the propensity of VT and NVT serotypes to cause illness as a result of carriage, by age, using estimates (and their uncertainty) from Choi et al. [24].

The impact of PCV on IPD was calculated based on VT and NVT disease propensity, as well as the vaccine efficacy against progression to invasive disease as a result of carriage (VE_{inv}). The latter was obtained from a function linking vaccine efficacy against IPD (VE_{IPD}) with VE_C and VE_{inv}, where $VE_{IPD} = 1 – (1- VE_c) *(1-VE_{inv})$ [25]. VE_{INV} was calculated based on estimates of VE_{IPD} from Lucero et al. [26] and estimates of VE_C from le Polain et al. [17] (Table1), as detailed in the Supporting Information (Part Two). Supporting Information (Part Two) also provides the details of the posterior estimates of the transmission parameters (i.e. the probability of transmission given contact) for VT and NVT, by age group, obtained in the fitting process.

Given the limited data on IPD available from Nha Trang and Vietnam [27], our main analysis focused on the proportion reduction in IPD, rather than IPD incidence. However, we illustrated how the relative IPD change may translate into disease incidence based on a point estimate of IPD in <5y olds of 49 /100,000 as reported for Nha Trang [27]. We did not infer disease impact among ≥5 year olds given a lack of data.

To explore the impact of delayed vaccination, we assumed that for each additional week of delay the incidence of IPD during that time would be stable at its pre-PCV level.

6.2.5. Vaccination strategies

We explored four different strategies: (i) routine vaccination only (RV), with two infant doses and a booster dose at 12 months of age ('2+1'), and routine vaccination with a catch-up campaign in (ii) <1y olds (CC1), (iii) <2y olds (CC2), and (iv) <5y olds (CC5).

WHO currently recommends introducing PCV either as a three primary infant dose schedule (3+0) or as two primary doses with a booster at 12 months of age (2+1 schedule), with the choice between schedules guided by setting-specific epidemiological characteristics [7]. We here present the predicted impact of a 2+1 programme, given the relatively low prevalence of carriage in young children in Nha Trang [27].

6.2.6. Sensitivity analyses

We ran the model for coverage levels of 50% and 70% in both routine and catch-up programmes, and compared the results with that of the main model (coverage 90%). We also explored the impact of duration of protection on our model outputs, based on lower values of 3 years and 20 years, which span across the range of likely values (Supporting Information Part Three).

6.2.7. Ethics

Ethical approval for both the contact and nasopharyngeal surveys was granted by the institutional review board of the London School of Hygiene and Tropical Medicine and the

Ethics Commission of National Institute of Hygiene and Epidemiology, Hanoi and the Nagasaki University.

6.3.   Results

The carriage prevalence in <5 year olds was 41% (95% CI 38 - 46%) overall, 27% (95%CI 23 - 32%) for VT serotypes and 14% (95%CI 11 - 18%) for NVT serotypes. We estimated the carriage prevalence of VT in 5 – 17 year olds and in adults to be 14% (95% CrI 10 – 18%) and 3% (95%CrI 0 – 7%) respectively, and that of NVT to be 15% (95%CrI 11 – 19%) and 3% (95%CrI 0 – 7%). The model fit to the carriage data is shown in Figure 20.

Figure 20: Pre-PCV carriage estimates across age groups for VT (left panel) and NVT (right panel), based on survey (plain vertical lines) and meta-regression model (dotted vertical lines) estimates, and estimates from the corresponding transmission model.
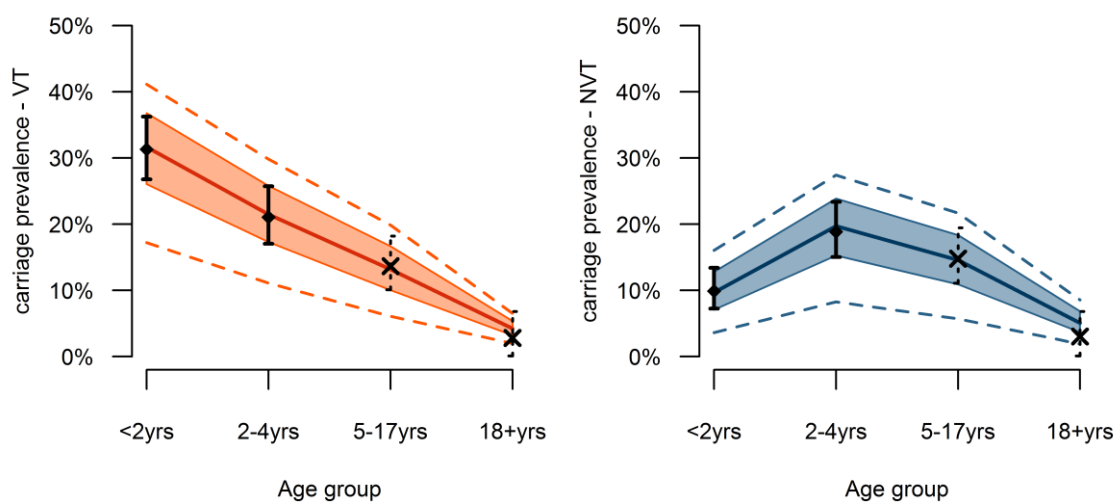
Legend: Dots and bars correspond to the point and 95% confidence interval for the carriage estimates from the survey data (plain dot and plain lines) and the meta-regression model (cross and dotted lines). The dark red and dark blue plain lines represent the median transmission model estimate for VT and NVT respectively, the shaded areas the 50% credible interval (CrI) around the median and the dotted red and blue lines the 95%CrI

6.3.1. Nasopharyngeal carriage

Our model predicts elimination of VT serotypes across all age groups within 10 years of PCV introduction in RV, with near-complete replacement by NVT serotypes, resulting in little or no change in the overall carriage prevalence (Figure 21A&B), particularly in children aged >5 years and in adults. CCs are predicted to reduce VT carriage more quickly through combined direct and indirect (i.e. herd) effects (Figures 21C&D).

Figure 21: Predicted trends in nasopharyngeal carriage following PCV13 introduction in Nha Trang.
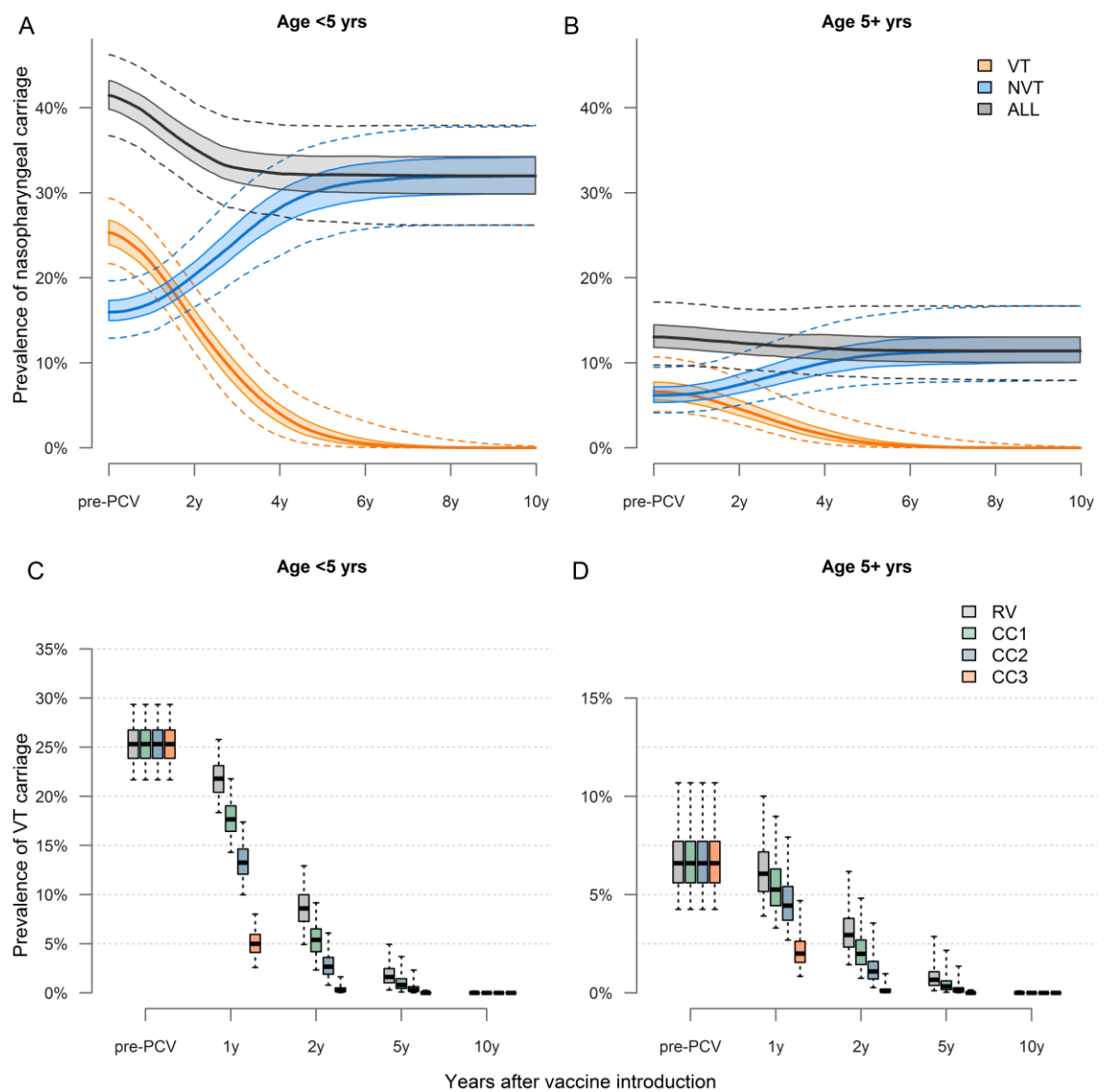
Legend Figure 1: A: Predicted trends in VT, NVT and overall carriage in <5 year olds without a catch-up campaign. B: Predicted trends in VT, NVT, and overall carriage in ≥5 year olds in RV. C: Predicted prevalence of VT carriage in <5 year

olds for each vaccination strategy. D: Predicted prevalence of VT carriage in ≥5 year olds. In all four panels: plain line= median, shaded areas= 50% CrI and dotted line or whiskers= 95% CrI

In children under five years the VT carriage in children is predicted to decrease by >99% within 6 years and 7 months (95% CI 5y2m - 9 y6m) in RV, 5 years and 11 months (4y6m– 8 y10m) in CC1, 5 years and 1 month (3y9m – 7y10m) in CC2 and 3 yrs (2y2m–4y10m) in CC5. Similar trends are predicted in older children and adults (Figures 21C and 21D). And in all scenarios NVT would increase proportionally to the decrease in VT.

6.3.2.   Invasive Pneumococcal Disease (IPD)

Our model predicts that the decline of IPD incidence will be proportionally highest among <2 year olds, falling to about 45% (95%CrI 33% - 57%) of its pre-PCV level and would be almost halved (57% (48 – 66%)) in children aged 2 – 4 years respectively (Figure 22A and 22B).

Most of the benefit of catch-up campaigns over routine vaccination in children <5 years is predicted to occur within the first three years after PCV introduction, with no noticeable difference at 5 years (Figure 22C). Our model predicts that the relative risk of IPD after CCs compared to RV would be lowest about 1 year after PCV introduction, with differences between strategies then steadily reducing thereafter until the new equilibrium is reached. Compared to RV, one year after vaccine introduction the number of cases of IPD is predicted to be 11% (95%CrI 9 – 14%) lower with CC1, 25% (21-30%) lower with CC2 and 38% (32 – 46%) lower with CC5 (Fig 22C).

The impact of each strategy on the cumulative proportion of cases averted in the first 3 years post PCV introduction, compared to no vaccination, is illustrated in Figure 22D.

Figure 22: Trends in IPD following PCV introduction in children under five years of age in Nha Trang
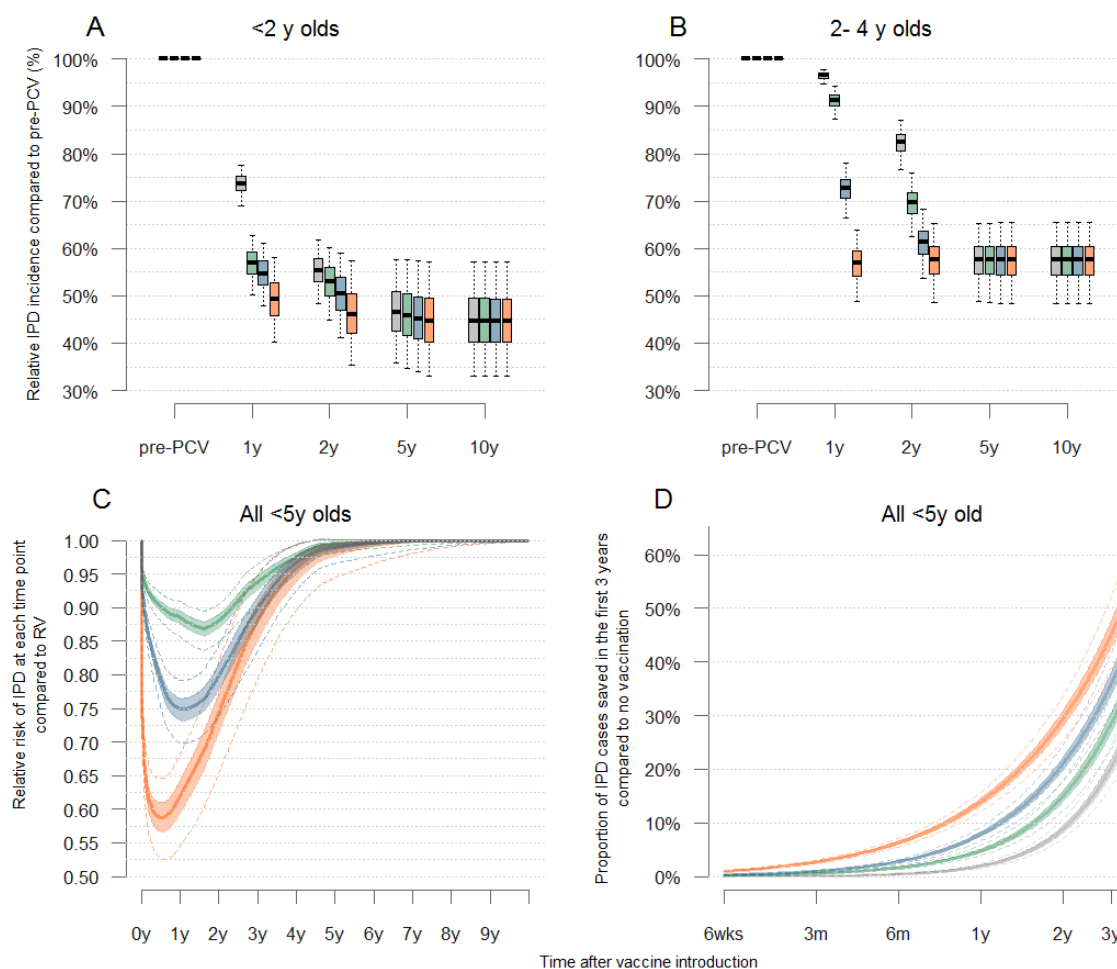
Legend: A: Predicted trends the cumulative annual incidence of IPD in children <2 years for each vaccination strategy considered, at a 90% vaccination coverage B: Predicted trends the cumulative annual incidence of IPD in children aged 2-4 years for each vaccination strategy considered, at a 90% vaccination coverage C: Cumulative number of IPD cases saved for each catch up strategy compared to RV, in children <5 years of age. D: Overall cumulative number of cases saved for each vaccination strategy, compared to no vaccination, in the first 3 years post PCV introduction. In all four panels: plain line= median, shaded areas= 50% credible intervals and dotted line or whiskers= 95% credible interval

Based on an average annual incidence risk of 49 cases per 100,000 children <5 years before PCV introduction [27], a routine introduction of PCV would result in a total of 74 cases (95%CrI 62 - 86) per 100,000 children <5 years averted over the first five years of programme implementation, considering all IPD (both VT and NVT), and catch-up campaigns would lead to the prevention of an additional 13 (95%CrI 11 – 16) cases with CC1, 25 (95%CrI 21 – 30) cases with CC2 and 39 (95%CrI 31 – 49) cases with CC5.

6.3.3. Delayed PCV introduction

We estimated the relative impact on IPD of catch-up campaigns for increasing delays. Our results suggest that, compared to RV, more IPD cases would be prevented in children <5 years insofar as PCV introduction is not delayed by more than 31 weeks (95%CI 30 – 32 weeks) for CC1, 58 weeks (53 – 63 weeks) for CC2 and 89 weeks (78 – 101 weeks) for CC5. Vaccination delays would negatively impact under 2 year olds more rapidly than 2 – 4 year old (Figure 23).

6.3.4. Sensitivity analyses

6.3.4.1.   Vaccination coverage

We explored model outputs with lower vaccination coverage in both cohort and catch-up immunization. Our model predicts a lengthening of the time to near-elimination of VT serotypes (and hence, the time reach the new post-PCV disease equilibrium) as vaccination coverage lowers but a similar differential impact of catch-up campaigns compared to routine vaccination. Full details are provided in the Supporting Information Part Three.

6.3.4.2.   Duration of protection

A duration of protection of 3 years would increase the time to elimination of VT carriage, and thus prevent fewer IPD cases overall, while any average duration of protection longer than 6 years would not change model outcomes. With a duration of 3 years the median prevalence of VT in <5 year olds is predicted to reach near elimination about 2 years later in RV and CC1, 1.5 years later with CC2 and about 1 year later with CC5. Similar differences were predicted for the ≥5 year olds (Figure S13 in the Supporting Information Part Three). The relative impact of one vaccination strategy over another was predicted to be similar than with a duration of 6 years.

Figure 23: Impact of delayed PCV introduction with a catch-up campaign on VT carriage (panel A) and on IPD cases saved (panel B) in children under five years of age.
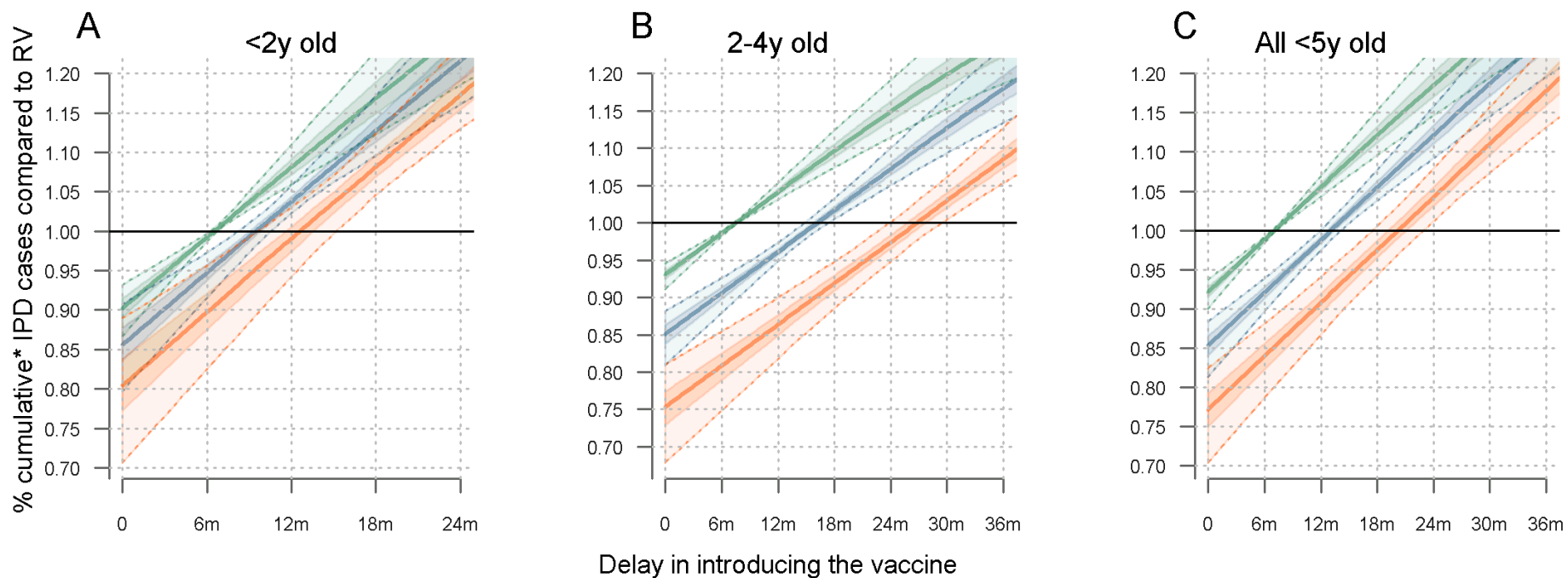
Legend: A: <2 year olds. B: 2 – 4 year olds C: <5 years olds. The middle plain trend line corresponds to the median estimate (Green=CC1, Blue=CC2, Red=CC5), the dark shaded areas the 50% CrI and the light shaded areas the 95%CrI. The plain horizontal line at 1.00 represents the point below which interventions will be more favourable than a timely RV.

6.4.   Discussion

We explored the possible impact of introducing PCV13 with and without a catch-up campaign in Vietnam through a dynamic transmission model. Our results feed into current debates about introduction strategies, particularly in South-East Asia where pneumococcal disease burden is high [28], where many countries have not yet introduced PCV [5], and where epidemiological data to guide decision making remain scarce [28, 29]. Although Vietnam is Gavi-eligible and is expected to introduce PCV in the coming years, it has not yet been approved for introduction [5]. Our results provide estimates about how much catch—up campaigns would decrease disease burden compared to routine vaccine introduction without a campaign, for different scenarios. Our study also shows that although catch-up campaigns would decrease disease burden more rapidly across age groups, their impact would only be beneficial insofar as the additional supply and operational constraints of their implementation does not delay PCV introduction by more than about 6 months to 2 years, depending on the age cohorts targeted by those campaigns.

The availability of data on both social mixing patterns – which are central to transmission models [13,30, 31] – and carriage in the same population allowed for thorough parameterisation of a transmission model.

However, our study has a number of limitations. In the absence of post-PCV data, predictions were based on the best available estimates of parameters governing vaccine effects that were observed elsewhere [16, 17, 32], which may not fully capture the local characteristics. Given that serotypes differ in their pathogenicity, fitness and transmissibility [33, 34], and that vaccine efficacy and duration of protection differs by serotype [12], our predictions based on

homogeneous characteristics for the group of VT and NVT serotypes may overlook local epidemiological characteristics. This uncertainty was nonetheless captured to some extent by sampling from the known uncertainty around those parameters [16]. Moreover, we were not able to assess the impact of PCV on pneumococcal pneumonia, the burden of which is much higher than that of IPD [35], given the lack of robust data and the challenges in the aetiological assessment of clinical pneumonia. Results from several large pneumonia aetiological studies [36] might help modelling work on the impact of PCV on pneumonia in the future.

Our predictions are in line with the experience of PCV7 in Europe and North America [37-44], as well with post-PCV trends observed in the few studies from resource-poor settings [45, 46]. What Kilifi has shown.

The implementation of PCV in various settings has consistently shown little or no change in overall carriage prevalence, due to replacement effects by NVT serotypes colonising the space left vacant by VT in the nasopharynx, but a reduction in severe disease given the lower pathogenicity of the latter, in accordance with our model output. Our predictions were also robust to estimates of duration of protection of vaccination coverage, and thus provide useful estimates of the impact of introducing PCV in a semi-urban Southeast Asian setting.

In situations where the number of vaccine doses available is limited, and ignoring any other supply side, staffing and outreach challenges that could potentially delay the implementation of CCs, our study can inform whether delaying the introduction of the vaccine to allow for CCs would potentially be beneficial, compared to a routine-only strategy. Moreover, this model also provides a framework that could feed into economic evaluations to further guide decisions about vaccine introduction with and without campaigns.

The generalisability to other settings of the differential impact of CCs needs to be considered in light of the epidemiological and socio-demographic characteristics of Nha Trang. In particular, the low prevalence of carriage and an ageing population - with only 5% of children under the age of five years – likely contribute to the rapid predicted reduction of disease due to PCV under all scenarios. Although similar epidemiological and demographic characteristics are observed in many other South(east) Asian settings [47, 48], the differential impact of CCs and the establishment of herd effects in settings with a younger population and a higher carriage prevalence is likely to differ, and should be addressed with models applied to such settings.

In conclusion, our study offers insights into the current debate about vaccination strategies when introducing PCV in South-East Asia. Our model suggests that catch-up campaigns have the potential to rapidly decrease carriage and disease across age groups, but are only offering added reduction in disease burden insofar their implementation results in little to no implementation delay.

Conflict of interest

OLP: none

WJE: WJE's partner works for GSK, who manufacture PCV10.

KT: none

KA: none

EKM: Kim Mulholland has consulted for GSK on PCV vaccine use and nutritional strategies to improve vaccine effectiveness.

DG: has served on ad-hoc advisory boards for Pfizer, GlaxoSmithKline and Merck, and the University College London Institute of Child Health Laboratory receives contract research funding from Pfizer, GlaxoSmithKline and Merck

References

[1] Cohen AL, Hyde TB, Verani J, Watkins M. Integrating pneumonia prevention and treatment interventions with immunization services in resource-poor countries. Bulletin of the World Health Organization. 2012;90:289-94.

[2] Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375:1969-87.

[3] Global Action Plan for prevention and control of pneumonia (GAPP). Geneva: World Health Organization; 2008.

[4] Calix JJ, Porambo RJ, Brady AM, Larson TR, Yother J, Abeygunwardana C, et al. Biochemical, genetic, and serological characterization of two capsule subtypes among Streptococcus pneumoniae Serotype 20 strains: discovery of a new pneumococcal serotype. J Biol Chem. 2012;287:27885-94.

[5] Gavi, The Vaccine Alliance. Advance Market Commitment for Pneumococcal Vaccines. Annual Report 1 April 2014 - 31 March 2015. In: Gavi TVA, editor.2015.

[6] Harris JB, Gacic-Dobo M, Eggers R, Brown DW, Sodha SV, Centers for Disease C, et al. Global routine vaccination coverage, 2013. MMWR Morbidity and mortality weekly report. 2014;63:1055-8.

[7] Publication WHO. Pneumococcal vaccines WHO position paper - 2012 - recommendations. Vaccine. 2012;30:4717-8.

[8] Whitney CG, Goldblatt D, O'Brien KL. Dosing schedules for pneumococcal conjugate vaccine: considerations for policy makers. The Pediatric infectious disease journal. 2014;33 Suppl 2:S172-81.

[9] Vu HT, Yoshida LM, Suzuki M, Nguyen HA, Nguyen CD, Nguyen AT, et al. Association between nasopharyngeal load of Streptococcus pneumoniae, viral coinfection, and radiologically confirmed pneumonia in Vietnamese children. The Pediatric infectious disease journal. 2011;30:11-8.

[10] Yoshida LM, Nguyen HA, Watanabe K, Le MN, Nguyen AT, Vu HT, et al. Incidence of radiologically-confirmed pneumonia and Haemophilus influenzae type b carriage before Haemophilus influenzae type b conjugate vaccine introduction in Central Vietnam. The Journal of pediatrics. 2013;163:S38-43.

[11] O'Brien KL, Nohynek H, World Health Organization Pneumococcal Vaccine Trials Carriage Working G. Report from a WHO Working Group: standard method for detecting upper respiratory carriage of Streptococcus pneumoniae. The Pediatric infectious disease journal. 2003;22:e1-11.

[12] Le Polain de Waroux O, Flasche S, Prieto-Merino D, Edmunds WJ. Age-dependent prevalence of nasopharyngeal carriage of streptococcus pneumoniae before conjugate vaccine introduction: a prediction model based on a meta-analysis. PloS one. 2014;9:e86136.

[13] Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoSMed. 2008;5:e74.

[14] Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ. What types of contacts are important for the spread of infections?: using contact survey data to explore European mixing patterns. Epidemics. 2011;3:143-51.

[15] Kadioglu A, Weiser JN, Paton JC, Andrew PW. The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. Nature reviews Microbiology. 2008;6:288-301.

[16] Choi YH, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, et al. 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PLoSOne. 2011;6:e26190.

[17] Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. The Pediatric infectious disease journal. 2015;34:858-64.

[18] Dagan R, Givon-Lavi N, Porat N, Greenberg D. The effect of an alternative reduced-dose infant schedule and a second year catch-up schedule with 7-valent pneumococcal conjugate vaccine on pneumococcal carriage: a randomized controlled trial. Vaccine. 2012;30:5132-40.

[19] Russell FM, Carapetis JR, Satzke C, Tikoduadua L, Waqatakirewa L, Chandra R, et al. Pneumococcal nasopharyngeal carriage following reduced doses of a 7-valent pneumococcal conjugate vaccine and a 23-valent pneumococcal polysaccharide vaccine booster. Clinical and vaccine immunology : CVI. 2010;17:1970-6.

[20] van Gils EJ, Veenhoven RH, Hak E, Rodenburg GD, Bogaert D, Ijzerman EP, et al. Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. JAMA : the journal of the American Medical Association. 2009;302:159-67.

[21] Ferraro CF, Trotter CL, Nascimento MC, Jusot JF, Omotara BA, Hodgson A, et al. Household crowding, social mixing patterns and respiratory symptoms in seven countries of the African meningitis belt. PloS one. 2014;9:e101129.

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

[23] Ali M, Canh GD, Clemens JD, Park JK, von Seidlein L, Minh TT, et al. The use of a computerized database to monitor vaccine safety in Viet Nam. Bulletin of the World Health Organization. 2005;83:604-10.

[24] Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. PLoSOne. 2012;7:e39927.

[25] Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL. The fundamental link between pneumococcal carriage and disease. Expert review of vaccines. 2012;11:841-55.

[26] Lucero MG, Dulalia VE, Nillos LT, Williams G, Parreno RA, Nohynek H, et al. Pneumococcal conjugate vaccines for preventing vaccine-type invasive pneumococcal disease and X-ray defined pneumonia in children less than two years of age. Cochrane database of systematic reviews. 2009:CD004977.

[27] Anh DD, Kilgore PE, Slack MP, Nyambat B, Tho le H, Yoshida LM, et al. Surveillance of pneumococcal-associated disease among hospitalized children in Khanh Hoa Province, Vietnam. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009;48 Suppl 2:S57-64.

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

[29] Jaiswal N, Singh M, Das RR, Jindal I, Agarwal A, Thumburu KK, et al. Distribution of serotypes, vaccine coverage, and antimicrobial susceptibility pattern of Streptococcus pneumoniae in children living in SAARC countries: a systematic review. PloS one. 2014;9:e108617.

[30] Edmunds WJ, O'Callaghan CJ, Nokes DJ. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proceedings Biological sciences / The Royal Society. 1997;264:949-57.

[31] Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. AmJEpidemiol. 2006;164:936-44.

[32] Auranen K, Mehtala J, Tanskanen A, Kaltoft S. Between-strain competition in acquisition and clearance of pneumococcal carriage--epidemiologic evidence from a longitudinal study of day-care children. AmJEpidemiol. 2010;171:169-76.

[33] Lipsitch M, Abdullahi O, D'Amour A, Xie W, Weinberger DM, Tchetgen Tchetgen E, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. Epidemiology. 2012;23:510-9.

[34] Hogberg L, Geli P, Ringberg H, Melander E, Lipsitch M, Ekdahl K. Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. Journal of clinical microbiology. 2007;45:948-52.

[35] Rudan I, El Arifeen S, Bhutta ZA, Black RE, Brooks A, Chan KY, et al. Setting research priorities to reduce global mortality from childhood pneumonia by 2015. PLoS medicine. 2011;8:e1001099.

[36] Levine OS, Bhat N, Crawley J, Deloria-Knoll M, DeLuca AN, Driscoll AJ, et al. Pneumonia etiology research for child health. Introduction. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2012;54 Suppl 2:S87-8.

[37] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoSMed. 2011;8:e1001017.

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

[39] Spijkerman J, van Gils EJ, Veenhoven RH, Hak E, Yzerman EP, van der Ende A, et al. Carriage of Streptococcus pneumoniae 3 years after start of vaccination program, the Netherlands. EmergInfectDis. 2011;17:584-91.

[40] Ghaffar F, Barton T, Lozano J, Muniz LS, Hicks P, Gan V, et al. Effect of the 7-valent pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae in the first 2 years of life. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2004;39:930-8.

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

[42] Hennessy TW, Singleton RJ, Bulkow LR, Bruden DL, Hurlburt DA, Parks D, et al. Impact of heptavalent pneumococcal conjugate vaccine on invasive disease, antimicrobial resistance and colonization in Alaska Natives: progress towards elimination of a health disparity. Vaccine. 2005;23:5464-73.

[43] Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, Reasonover A, et al. Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. JInfectDis. 2006;193:1487-94.

[44] Nzenze SA, Shiri T, Nunes MC, Klugman KP, Kahn K, Twine R, et al. Temporal changes in pneumococcal colonization in a rural African community with high HIV prevalence following routine infant pneumococcal immunization. The Pediatric infectious disease journal. 2013;32:1270-8.

[45] Roca A, Hill PC, Townend J, Egere U, Antonio M, Bojang A, et al. Effects of community-wide vaccination with PCV-7 on pneumococcal nasopharyngeal carriage in the Gambia: a cluster-randomized trial. PLoSMed. 2011;8:e1001107.

[46] Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. The Lancet Global health. 2014;2:e397-405.

[47] Adegbola RA, DeAntonio R, Hill PC, Roca A, Usuf E, Hoet B, et al. Carriage of Streptococcus pneumoniae and other respiratory bacterial pathogens in low and lower-middle income countries: a systematic review and meta-analysis. PloS one. 2014;9:e103293.

[48] UNFPA. State of the World Population 2014. 2015.

[49] Geweke, *Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments*, in *Bayesian Statistics* J.B. J.M. Bernardo, A.P. Dawid and A.F.M. Smith, Editor. 1992, Oxford University Press: Oxford. p. 169 - 193.

6.5. Supporting Information

6.5.1. Part 1: Model Structure

We built a realistic age-structured deterministic Susceptible-Infected-Susceptible (SIS) transmission model of carriage acquisition and clearance, similar to Choi et al. [1].

The model considers all VT serotypes jointly and separately from the group of NVT serotypes, but allows for VT-NVT co-colonization, thus resulting in four compartments of carriage states, namely susceptible, VT carriers, NVT carriers and VT-NVT co-colonized as in Choi et al. [1]. In this model movements from the susceptible to the VT or NVT compartments are determined by age-specific forces of infection for VT (λ_{Vi}) and NVT (λ_{Ni}) respectively, and co-colonization is determined by competition parameters (C_N and C_V), which represent the degree with which prior colonization reduces the likelihood of co-colonization. Age-specific recovery rates (r_i) determine the speed with which individuals revert back from the co-colonized to either VT or NVT compartments, and from colonization to susceptible compartments, assuming no natural immunity to carriage.

The model comprised of three levels of vaccine-induced immunity; (1) no protection, (2) partial protection and (3) full protection (Figure S9). The latter means the efficacy and duration of protection conferred after completion of the infant schedule (i.e. 2 infant doses and a booster at 12 months ('2+1' schedule)) or the completion of a catch-up programme in older children (2 doses in <18 months and 1 dose in ≥18 months). Partial protection was gained form two primary infant doses, or after the first catch-up dose in children aged 12 - 17 months. The difference between full and partial protection lied in the magnitude of vaccine efficacy against carriage (VE_C) and in the duration of protection.

We applied the model to a population of 81 annual age cohorts (0 to 80 years) divided into 52 weekly age bands of 100 individuals. In the calculation of the force of infection the population figure was adjusted to represent the actual population, based on census data.

We inferred the impact on IPD based posterior estimates of carriage at each time step, case:carrier ratios for VT and NVT by age group [1], the vaccine efficacy against invasiveness and the vaccination coverage of each weekly age cohort at each time step. Details about parameters of vaccine efficacy are provided in the Supporting Information Part Two.

The model equations are provided in Text S1 below.

Figure S9: The three groups of vaccine protection states defined in the model and movement between groups
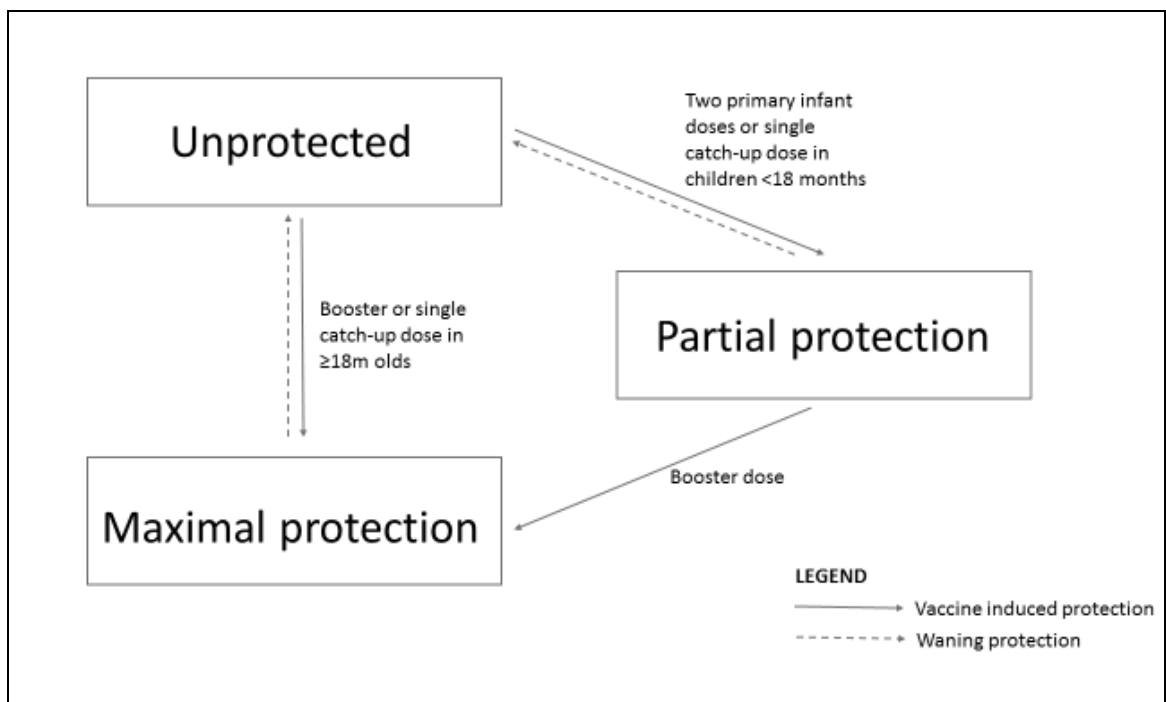
Text S1: Model equations

$$\frac{dS_i(t)}{dt} = r_{Vi} * V_i(t) + r_{Ni} * N_i(t) - (\lambda_{Vi}(t) + \lambda_{Ni}(t)) * (\gamma_{i(\text{inf})} + \gamma_{i(\text{catch} - up)}) * S_i(t) + \omega_p * Sp_i(t) + \omega_f * Sf_i(t)$$

$$\frac{dV_i(t)}{dt} = \lambda_{Vi}(t) * S_i(t) + r_{Ni} * B_i(t) - (C_N * \lambda_{Ni}(t) + r_{Vi}) * V_i(t) - \gamma_{i(\text{inf})} + \gamma_{i(\text{catch} - up)}) * V_i(t) + \omega_p * Vp_i(t) + \omega_f * Vf_i(t)$$

$$\frac{dN_i(t)}{dt} = \lambda_{Ni}(t) * S_i(t) + r_{Vi} * B_i(t) - (C_V * \lambda_{Vi}(t) + r_{Ni}) * N_i(t) - \gamma_{i(\text{inf})} + \gamma_{i(\text{catch-up})}) * N_i(t) + \omega_p * Np_i(t) + \omega_f * Nf_i(t)$$

$$\frac{dB_i(t)}{dt} = (C_V * \lambda_{Vi}(t) * N_i(t) + C_N * \lambda_{Ni}(t) * V_i(t) - (r_{Ni} + r_{Vi}) * B_i(t) - (\gamma_{i(\text{inf})} + \gamma_{i(\text{catch-up})}) * B_i(t) + \omega_p * Bp_i(t) + \omega_f * Bf_i(t)$$

$$\frac{dSp_i(t)}{dt} = r_{Vi} * Vp_i(t) + r_{Ni} * Np_i(t) - ((1 - VE_{pc}) * \lambda_{Vi}(t) + \lambda_{Ni}(t) + \gamma_{i(\text{booster})} + \omega_p) * Sp_i(t) + \gamma_{i(\text{inf})} * S_i(t)$$

$$\frac{dVp_i(t)}{dt} = (1 - VE_{pc}) * \lambda_{Vi}(t) * Sp_i(t) + r_{Ni} * Bp_i(t) - (C_N * \lambda_{Ni}(t) + \gamma_{i(\text{booster})} + \omega_p) * Vp_i(t) + \gamma_{i(\text{inf})} * V_i(t)$$

$$\frac{dNp_i(t)}{dt} = \lambda_{Ni}(t) * Sp_i(t) + r_{Vi} * Bp_i(t) - (C_V * (1 - VE_{pc}) * \lambda_{Ni}(t) + \gamma_{i(\text{booster})} + \omega_p) * Np_i(t) + \gamma_{i(\text{inf})} * N_i(t)$$

$$\frac{dBp_i(t)}{dt} = (C_V * (1 - VE_{pc}) * \lambda_{Vi}(t) * Np_i(t) + C_N * \lambda_{Ni}(t) * V_i(t) - (r_{Ni} + r_{Vi} + \gamma_{i(\text{booster})} + \omega_p) * Bp_i(t) + \gamma_{i(\text{inf})} * B_i(t)$$

$$\frac{dSf_i(t)}{dt} = r_{Vi} * Vf_i(t) + r_{Ni} * Nf_i(t) - ((1 - VE_{fc}) * \lambda_{Vi}(t) + \lambda_{Ni}(t) + \omega_f) * Sf_i(t) + \gamma_{i(\text{booster})} * Sp_i(t) + \gamma_{i(\text{catch-up})} * Sp_i(t)$$

$$\frac{dVf_i(t)}{dt} = (1 - VE_{fc}) * \lambda_{Vi}(t) * Sf_i(t) + r_{Ni} * Bf_i(t) - (C_N * \lambda_{Ni}(t) + \omega_f) * Vf_i(t) + \gamma_{i(\text{booster})} * Vp_i(t) + \gamma_{i(\text{catch-up})} * Vp_i(t)$$

$$\frac{dNf_i(t)}{dt} = \lambda_{Ni}(t) * Sf_i(t) + r_{Vi} * Bf_i(t) - (C_V * (1 - VE_{fc}) * \lambda_{Ni}(t) + \omega_f) * Nf_i(t) + \gamma_{i(\text{booster})} * Np_i(t) + \gamma_{i(\text{catch-up})} * Np_i(t)$$

$$\frac{dBf_i(t)}{dt} = (C_V * (1 - VE_{fc}) * \lambda_{Vi}(t) * Nf_i(t) + C_N * \lambda_{Ni}(t) * V_i(t) - (r_{Ni} + r_{Vi} + \omega_f) * Bf_i(t) + \gamma_{i(\text{booster})} * Bp_i(t) + \gamma_{i(\text{catch-up})} * Bp_i(t)$$

- i represent the weekly age groups.

- Υ is the vaccination coverage for infant doses (inf), the booster dose (boost) and the catch-up dose (catch-up).

- VE_{fc} and VE_{pc} is the vaccine efficacy against carriage acquisition after full and partial vaccination respectively.

- ω_p and ω_f represent the rate of waning immunity after partial and full vaccination

6.5.2. Part Two: Estimates of vaccine efficacy, and transmission probabilities

6.5.2.1. Vaccine efficacy against carriage

Full vaccination. We obtained the vaccine efficacy against carriage after full protection (VE_C^F) and its uncertainty from a Bayesian meta-regression model based on estimates from 22 intervention studies. The details have been published elsewhere [2]. The model of vaccine efficacy was run for as many iterations as in our transmission model, so that we could sample one estimate of VE_C^F at each iteration of the post-vaccination simulations.

Partial vaccination. A systematic review of the impact of pneumococcal conjugate vaccines on carriage [3] provided information on intervention studies with 2 primary-dose arms where compared with either 2+1 schedules or 3+0 schedules.

We included individual randomized controlled trials providing nasopharyngeal carriage estimates within 3 – 20 months after a 2-dose infant schedule (2+0 arm) and after a complete schedule (either 3+0 or 2+1 arm). We excluded studies providing carriage estimates within 3 months after vaccination to account for the delay in producing an immune response, and time to carriage clearance under the assumption that existing carriage at the time of vaccination is not affected by PCV [2].

We extracted data on the number of carriers of VT and total number of individuals swabbed in each study arm before and after vaccination. We then computed the relative risk of VT carriage among fully vaccinated compared to partially vaccinated children. The pooled estimates were obtained through a random effects meta-regression model.

We pooled estimates from four studies, including trials in Fiji [4], Israel [5], The Gambia [6] and The Netherlands [7], providing together eight different survey points, into a simple random effects meta-analysis, not taking into account within-study dependence given the limited number of data points per study.   The details of the studies are provided in Table S6.

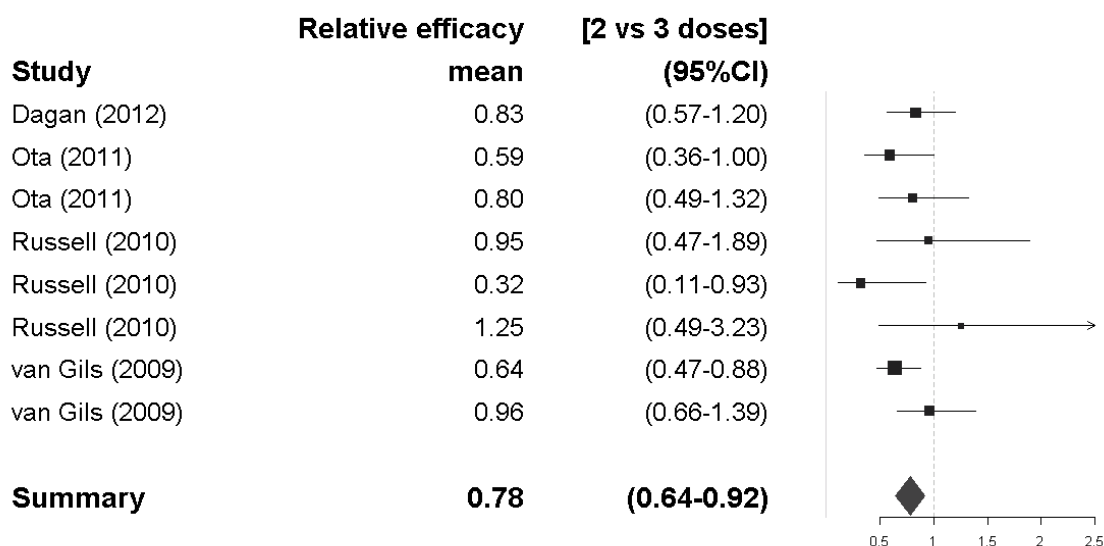Table S6: Studies included in the meta-analysis

Study	Country	Schedule	Age (m**)	Time since last dose (m)	Carriers /total	RR (95%I)
Dagan (2012)[5]*	Israel	3+0	9.5m	3.5m	58/327	0.83 (0.57 – 1.20)
Dagan (2012)[5]*	Israel	2+0	9.5m	3.5m	36/169	ref
Ota (2011)[6]	Gambia	3+0	11m	7m	20/200	0.59 (0.36 – 1.00)
Ota (2011)[6]	Gambia	2+0	11m	8m	33/198	ref
Ota (2011)[6]	Gambia	3+0	15m	11m	24/193	0.80 (0.49 – 1.32)
Ota (2011)[6]	Gambia	2+0	15m	12m	30/196	ref
Russell (2010)[4]	Fiji	3+0	6m	3m	13/127	0.95 (0.47 – 1.89)
Russell (2010)[4]	Fiji	2+0	6m	3m	16/148	ref
Russell (2010)[4]	Fiji	3+0	9m	6m	4/122	0.32 (0.11 – 0.93)
Russell (2010)[4]	Fiji	2+0	9m	6m	15/146	ref
Russell (2010)[4]	Fiji	3+0	12m	9m	8/114	1.25 (0.49 – 3.23)
Russell (2010)[4]	Fiji	2+0	12m	9m	9/143	ref
van Gils (2009)[7]	Netherlands	2+1	18m	7m	51/329	0.64 (0.47 – 0.88)
van Gils (2009)[7]	Netherlands	2+0	18m	14m	79/327	ref
van Gils (2009)[7]	Netherlands	2+1	24m	13m	47/333	0.96 (0.66 -1.39)
van Gils (2009)[7]	Netherlands	2+0	24m	20m	49/332	ref

*In this study pooled results were provided for samples taken at 7 months and 12 months of age, and we therefore considered the age and time since the last dose as the average between the two, **m=months

The pooled relative risk of carriage among children with a full schedule was 78% (95%CI 64 – 92%) that of a partial schedule (Figure S10), with no evidence of heterogeneity (I^2=3%). At each iteration of the MCMC, we sampled a value from the relative risk 78% (95%CI 64 – 92%) multiplied by a value of the full vaccine efficacy against carriage acquisition (VE_C^F), to obtain a measure of the partial efficacy (VE_C^P).

Figure S10 : Forest plot of the relative risk of VT carriage of 3 primary doses (2+1 or 3+0), for four different individual randomized controlled trials (RCTs), at eight different time points [4-7].

Study	Relative efficacy mean	[2 vs 3 doses] (95%CI)	
Dagan (2012)	0.83	(0.57-1.20)	
Ota (2011)	0.59	(0.36-1.00)	
Ota (2011)	0.80	(0.49-1.32)	
Russell (2010)	0.95	(0.47-1.89)	
Russell (2010)	0.32	(0.11-0.93)	
Russell (2010)	1.25	(0.49-3.23)	
van Gils (2009)	0.64	(0.47-0.88)	
van Gils (2009)	0.96	(0.66-1.39)	
Summary	**0.78**	**(0.64-0.92)**	

6.5.2.2. *Duration of protection*

The duration of protection against carriage acquisition was obtained from le Polain et al [2]. The median estimate of a model of waning efficacy, using an asymptomatic function, closely matched that of an exponential decay with a mean duration of protection of 6 years (Figure S9). Hence, this estimate was taken as the average duration of protection in the main model. For sensitivity analyses, we considered mean durations of 3 years and 20 years as the lower and higher values, as shown in Figure S11.

In the absence of estimates on the duration of protection of partial vaccination, we assumed that the duration of protection of a partial vaccination was 0.78 (0.64 – 0.92) that of a full vaccination, as for VE_C.

Figure S11: Duration of protection of PCV against carriage acquisition
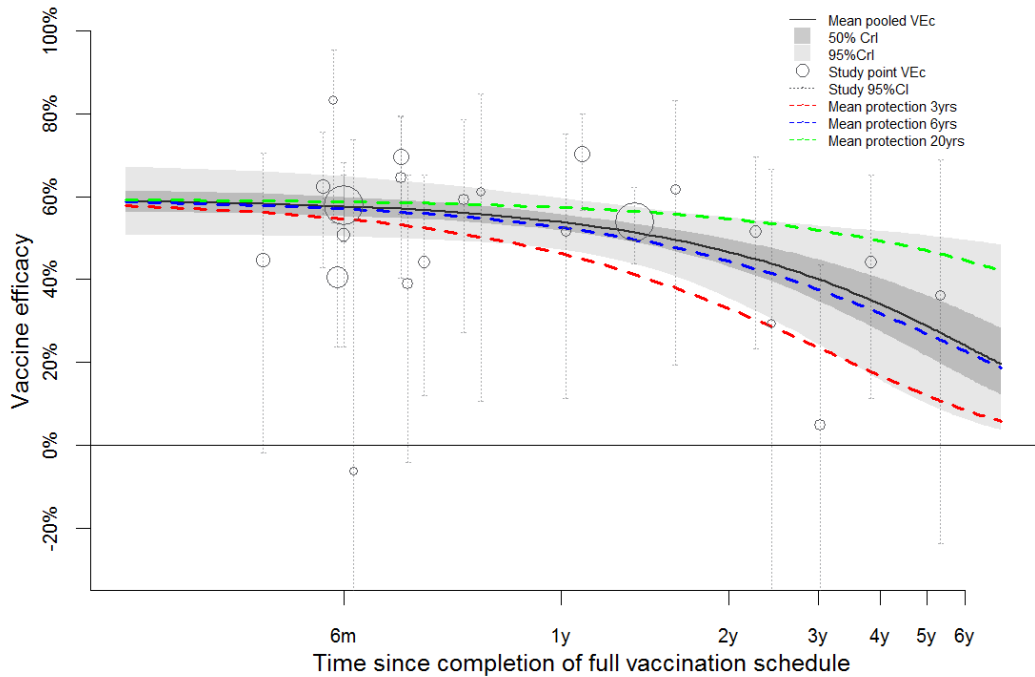
Legend : Each circle represents the mean vaccine efficacy estimate for each study included in the analysis – see [2]. The size of the circle is proportional to the study size. The whiskers on either side of each circle represent the 95%CrI around the point estimate. The dark grey area corresponds to the 50% CrI around the main model estimate of vaccine efficacy and its waning (plain black line), and the lighter grey area to the 95% CrI. The blue, red and green dotted lines show the waning with an exponential decay function for a mean duration of protection of 6 years, 3 years and 20 years respectively. Figure adapted from le Polain de Waroux et al. [2], with permission from PIDJ
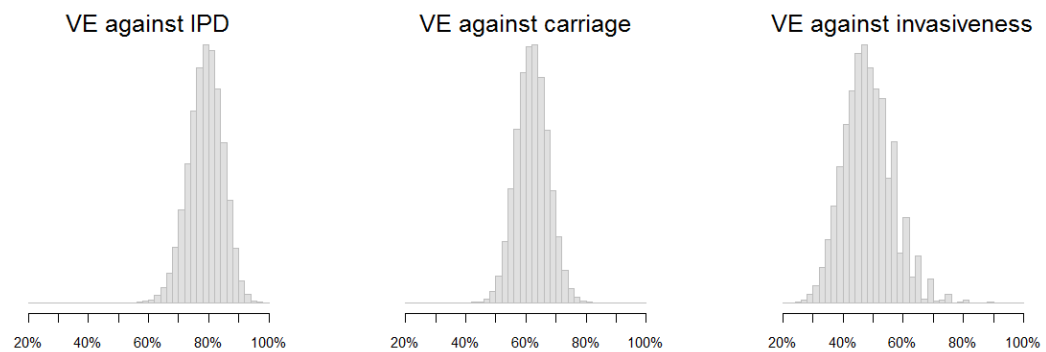
6.5.2.3. Vaccine efficacy against Invasive Pneumococcal Disease

The vaccine efficacy against invasive pneumococcal disease (VE_{IPD}) can be expressed as a function of the vaccine efficacy against carriage acquisition (VE_C) and the efficacy against progression to disease as a result of carriage, which we here term the vaccine efficacy against invasiveness (VE_{inv}) [8], where $VE_{IPD} = 1 - (1-VE_C)(1-VE_{inv})$.

Based on estimates from a large systematic review by Lucero et al. [9], we assumed VE_{IPD} to be 80% (95%CI 58 − 90%) and generated a binomial distribution that closely matched those estimates . Hence, we calculated VE_{inv} based on the distribution of VE_{IPD} and that of VE_C, and obtained a value of VE_{inv} at each iteration in our MCMC process.

The mean VE_{INV} was 80% (95%CrI 67 − 90%), the mean VE_C was 62% (95%CrI 52 − 73%) and mean VE_{inv} was 48% (95%CrI 34 − 65%). The distributions are shown in Figure S12.

Figure S12: Estimates of vaccine efficacy

Legend: density plot of the vaccine efficacy against IPD (left panel), carriage (medium panel) and invasiveness (right panel). The x axis displays the vaccine efficacy.

In the model, was assumed no waning of VE_{inv}, and that VE_{inv} would be conferred after 2 primary doses, or any catch-up dose. Although there is evidence that VE_C after partial vaccination differs from that after complete vaccination (see earlier), evidence that the efficacy against progression to invasiveness differs between two and three primary doses is scarce.

Although it is likely that VE_{inv} wanes over time, we did not consider waning in the main model, given that most of the direct impact of PCV on IPD occurs in the first few years of life, before

establishment of the herd immunity effect (see results), and given the lack of estimates of waning of VE$_{INV}$.

6.5.2.4.   *Transmission probabilities (posterior estimates)*

The probability of infection for NVT and VT by age group was obtain from fitting the model to the nasopharyngeal carriage data. The posterior estimates, including median and 95% credible intervals (CrI) are displayed in the Table S7 below

Table S7: Posteriors for the probability of infection by age group

Age group	<3y	3 – 5y	6 – 19y	20+y
Probability of infection with VT if contacted (10^{-2})	3.8 (2.8 – 5.2)	3.1 (2.2 – 5.7)	1.8 (1.5 – 2.0)	1.0 (0.7 – 1.4)
Probability of infection with NVT if contacted (10^{-2})	1.2 (0.7 – 1.9)	3.8 (2.6 – 4.9)	1.8 (1.4 – 2.1)	1.3 (0.9 – 1.8)

6.5.3. Part Three: Sensitivity analysis (coverage and duration of vaccine protection)

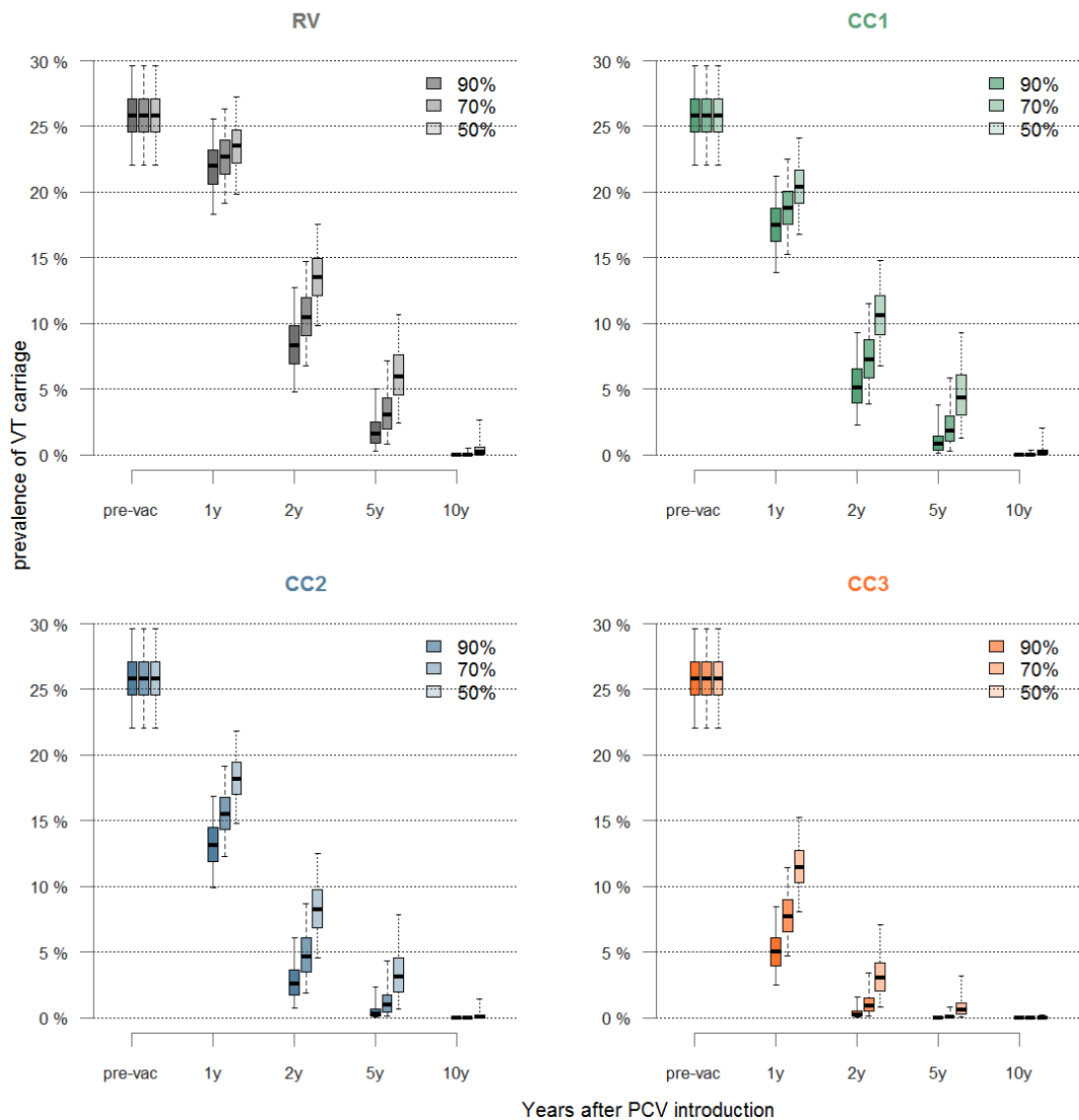*6.5.3.1.*   *Vaccination coverage*

We explored model outputs with lower vaccination coverage in both cohort and catch-up immunization. Our model predicted a lengthening of the time to near-elimination of VT serotypes (and hence, the time reach the new post-PCV disease equilibrium) as vaccination coverage lowers but a similar differential impact of catch-up campaigns compared to routine vaccination.

Figure S13 compares predictions of trends in VT carriage for all four strategies and vaccination coverage levels of 90%, 70% and 50%.
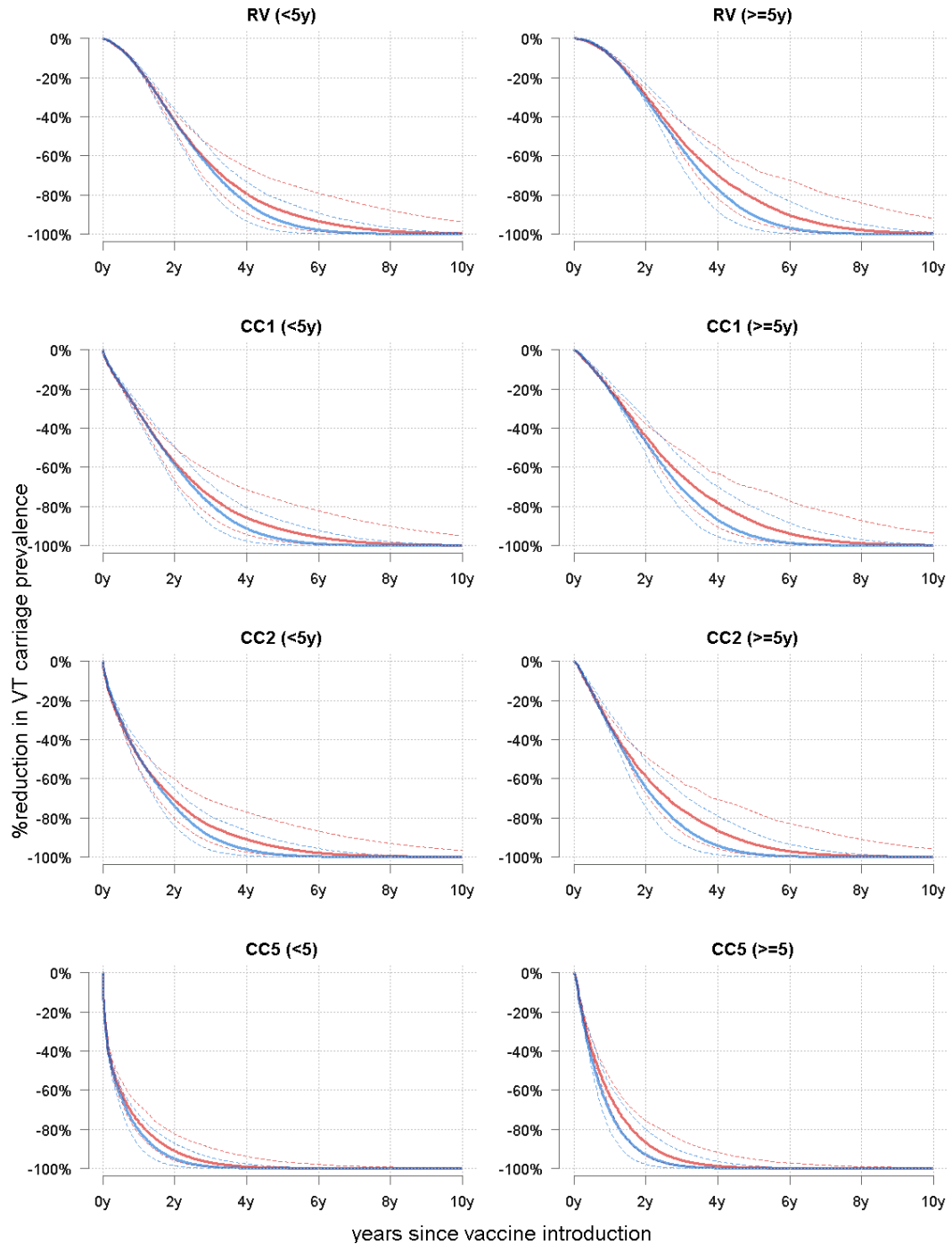
6.5.4. Duration of protection

A duration of protection of 3 years would increase the time to elimination of VT carriage, and thus prevent fewer IPD cases overall, while any average duration of protection longer than 6 years would not change model outcomes. With a duration of 3 years the median prevalence of VT in <5 year olds is predicted to reach near elimination about 2 years later in RV and CC1, 1.5 years later with CC2 and about 1 year later with CC5. Similar differences were predicted for the ≥5 year olds (Figure S14). The relative impact of one vaccination strategy over another was predicted to be similar than with a duration of 6 years.

Figure S13 Impact of vaccination coverage on VT carriage for each strategy considered, and comparing 90%, 70% and 50% coverage for each of the four strategies, in children aged 0 – 59 months

Legend: In all four panels: plain horizontal line= median, boxes= 50% credible intervals and whiskers= 95% credible interval, for each of the three levels of vaccination coverage (90%, 70%, 50%) and for all four vaccination strategies.

Figure S14: Comparing model estimates with an average duration of protection of 6 years (blue) and of 3 years (red) for each of the vaccination scenarios considered and by age group

Legend: Percentage reduction in VT carriage in children <5 years (left) and ≥ 5 years old. In blue is are estimates of the main model (average duration of protection of 6 years) and in red are estimates considering a mean duration of protection of 3 years. The plain lines represent the median and the dotted lines the 95% credible interval

References

[1] Choi YH, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, et al. 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PLoSOne. 2011;6:e26190.

[2] Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. The Pediatric infectious disease journal. 2015;34:858-64.

[3] Fleming-Dutra KE, Conklin L, Loo JD, Knoll MD, Park DE, Kirk J, et al. Systematic review of the effect of pneumococcal conjugate vaccine dosing schedules on vaccine-type nasopharyngeal carriage. The Pediatric infectious disease journal. 2014;33 Suppl 2:S152-60.

[4] Russell FM, Carapetis JR, Satzke C, Tikoduadua L, Waqatakirewa L, Chandra R, et al. Pneumococcal nasopharyngeal carriage following reduced doses of a 7-valent pneumococcal conjugate vaccine and a 23-valent pneumococcal polysaccharide vaccine booster. Clinical and vaccine immunology : CVI. 2010;17:1970-6.

[5] Dagan R, Givon-Lavi N, Porat N, Greenberg D. The effect of an alternative reduced-dose infant schedule and a second year catch-up schedule with 7-valent pneumococcal conjugate vaccine on pneumococcal carriage: a randomized controlled trial. Vaccine. 2012;30:5132-40.

[6] Ota MO, Akinsola A, Townend J, Antonio M, Enwere G, Nsekpong D, et al. The immunogenicity and impact on nasopharyngeal carriage of fewer doses of conjugate pneumococcal vaccine immunization schedule. Vaccine. 2011;29:2999-3007.

[7] van Gils EJ, Veenhoven RH, Hak E, Rodenburg GD, Bogaert D, Ijzerman EP, et al. Effect of reduced-dose schedules with 7-valent pneumococcal conjugate vaccine on nasopharyngeal pneumococcal carriage in children: a randomized controlled trial. JAMA : the journal of the American Medical Association. 2009;302:159-67.

[8] Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL. The fundamental link between pneumococcal carriage and disease. Expert review of vaccines. 2012;11:841-55.

[9] Lucero MG, Dulalia VE, Nillos LT, Williams G, Parreno RA, Nohynek H, et al. Pneumococcal conjugate vaccines for preventing vaccine-type invasive pneumococcal disease and X-ray defined pneumonia in children less than two years of age. Cochrane database of systematic reviews. 2009:CD004977.

7. Discussion

7.1. Main findings and strengths

The chapters of this thesis are presented in a sequence that reflects the steps that were taken to address the main questions around the epidemiology and transmission dynamics of pneumococci, and questions around PCV introduction strategies, which the thesis intended to address.

The first research paper of this thesis [1] provided further insight into the dynamics of nasopharyngeal carriage, showing a strong correlation between age groups in the prevalence of carriage of *Streptococcus pneumoniae* – i.e. that in populations where the carriage among children is high, carriage prevalence in adults is high too, and vice-versa. Results from *Research paper 1* [1] also suggest a 'shift' from VT serotypes – which are most commonly found in childhood [2] – to NVT serotypes, as age increases [3]. This may be explained by serotype-specific mucosal immunity acquired against some of the most common childhood serotypes, - predominantly VT serotypes - , which leads to colonisation by other serotypes instead. Under such hypothesis, a more balanced distribution between VT and NVT would occur earlier in life in high burden settings. Although *Research paper 1* was underpowered to address this question, a larger systematic collection of nasopharyngeal carriage data, which is currently underway (see subchapter 7.3), could make it possible to explore this further.

Research paper 1 [1] was designed to address a specific need, namely that of predicting the carriage prevalence, and distribution of VT and NVT serotypes in older children (5 – 17 years old) and adults (≥18 years), based on carriage estimates on children <5 years, in order to obtain precise estimates for Nha Trang, where data in older children and adults were not available.

However, the results have in fact wider implications, including for this thesis, as well as for further research, in addition to those that have already been discussed in the paper itself.

First, the statistical model was used to help sample size calculations for the nasopharyngeal carriage study conducted in Sheema district, Southwest Uganda (*Research papers 3 and 4*), based on limited data from another survey in children <5 years conducted in Uganda [4]. The nasopharyngeal carriage study, which is currently being written up (Nackers et al. in preparation and [5]), included over 1800 participants, from all age groups.

Second, in light of the increased emphasis on nasopharyngeal carriage as a proxy to monitor, predict and evaluate the population impact of PCV [6-9], as well as recent debates around pneumococcal vaccination strategies in older adults and elderly in countries where PCV programmes are well established [10-13], results from *Research paper 1* [1] will help better predicting carriage and relative prevalence of VT serotypes among older age groups, in settings where only data in children <5 years of age are available. This would also be useful in assessing the potential impact of vaccination strategies on invasive disease in older adults and in the elderly, based on estimates of invasiveness – and assuming those are stable [14].

Finally, *Research paper 1* also highlights important differences in the prevalence and vaccine-type serotype carriage between settings, with a carriage prevalence in under-fives as low as 21% in Taiwan and as high as 93% in The Gambia, and important differences in the carriage prevalence among adults. Stark variations in carriage prevalence estimates have also been recently highlighted in two systematic reviews of pneumococcal carriage prevalence [15, 16]. Drawing from the results from *Research paper 4*, it is plausible that differences in sociodemographic structures and household sizes, particularly the number of close contacts that individuals have on a daily basis, explain some of the patterns seen, and this would need further research. For example, the lower prevalence of nasopharyngeal carriage observed in

247

Nha Trang compared to Sheema district may be driven by differences in socio-demographic and social mixing patterns between these settings. Vietnam is currently experiencing a rapid demographic transition [17], with low mortality, low fertility rates and increasing life expectancy (average of 78 years in the latest census [17]). The average household size has declined in Vietnam from 5.1 in the late 1970s to around 3.7 in the 2009 census, largely due to a reduction in fertility rates over the last years, and a decrease in both the proportion and absolute number of children under five years of age in the community, who now represent just 8% of the total population [17]. In contrast, Uganda has one of the youngest populations, with 19% of the population aged <5 years, an average of 5.1 members per household, and a life expectancy of 50.4 years, as reported in the latest Demographic and Health Survey [18]. Drawing further parallels, it is interesting to notice that countries in West Africa, which have the highest prevalence of carriage (as high as >90% in children <5 years and >50% among adults in data from the Gambia) [15, 16] are also settings where household sizes are among the largest (the average household size in the Gambia is 8.2, and as high as 9.9 in rural settings [19]). In line with findings from *Research paper 4* about contacts relevant for carriage acquisition, and findings from *Research paper 1* of the adult to child proportionality in carriage across settings, such socio-demographic factors may potentially explain some of the epidemiological differences observed.

Pneumococcal transmission dynamics, population herd effects of PCV, and the differential impact of catch-up campaigns therefore likely differ in higher burden setting such as Sheema district, compared to the results displayed in the Nha Trang model, and further research is required to better understand how PCV impact would differ between such settings. This is further discussed in 7.3.

As for the first paper, the rationale for *Research paper 2* [20] was primarily to inform the modelling work applied to Nha Trang, and specifically compile estimates of vaccine efficacy

248

against carriage acquisition, and of the duration of protection. Previous estimates, in particular of the duration of carriage used in transmission models, were mostly based on anecdotal evidence [21]. Hence, the pooled evidence based on data from Europe, North America, South Africa and Israel, generated in *Research paper 2*, fills an important gap. The study provided the first ever estimates of the vaccine efficacy against carriage at different time points after vaccination, thereby improving our understanding of whether and how the protection against carriage wanes, and importantly, how this varies between serotypes and schedules.

The immediate application of *Research paper 2* was two-fold. First, given the lack of robust estimates on both vaccine efficacy and duration of protection, estimates directly fed into the transmission model (*Research paper 5*). Bayesian regression model fitting allowed using the full posterior estimates of vaccine efficacy in the transmission model. Efficacy estimates from *Research paper 2*, in combination with meta-estimates of the vaccine efficacy against IPD [22] were also used to obtain meta-estimates of the vaccine efficacy against invasiveness – for which there was no good estimate – using a formula described elsewhere [23], to estimate the impact of PCV on invasive disease in the Nha Trang model.

Second, results from *Research paper 2* also provided more robust estimates and more plausible uncertainty ranges, on the duration of protection, which resulted in more informed choices for sensitivity analyses.

The application of the results from *Research paper 2* to specific contexts will depend on the distribution of circulating serotypes, given the differences between serotypes highlighted in our analysis. The serotype distribution in the studies included in the paper [20] (n=15) was relatively homogeneous, and also consistent with the serotype distribution in Nha Trang [24]. Given that *Research paper 2* provided the first robust meta-estimates of differences in vaccince efficacy against carriage for individual serotypes and between schedules, future transmission models applied to settings where the serotype distribution differs substantially from the studies

included in *Research paper 2,* serotype-specific estimates of vaccine efficacy against carriage may need to be considered.

Findings from *Research papers 3* provided essential data on the social contact patterns in a rural African setting. The comparison of mixing patterns from Great Britain [25] and Uganda – showing more assortative mixing in Britain and more intergenerational mixing in Uganda -, and subsequent differences in simulations of the final epidemic size, emphasised the importance of applying setting-specific contact structured to transmission models. While *Research paper 3* did not directly contribute to other outputs from this thesis, it was novel and informative in various ways. First, given the scarcity of contact data from low-income settings, let alone Africa, the study filled an important gap by providing very detailed estimates of the structure of social contacts between individuals, by age and by type of encounter. It also offered valuable insights into the spatial characteristics of social encounters. Spatial interconnectedness has increasingly been recognized as an essential factor driving epidemic extinction or persistence of epidemic hotspots, and the effectiveness of control strategies. However, even though spatial components have been explored in details in some studies, such as Read et al. in China [26], information on the spatial dimension of social contacts has often been overlooked, and is certainly lacking for rural African settings. *Research paper 3* is the first study providing quantitative information on the local spatial aspect of contacts in a low-income African setting, which will be very valuable for infectious disease modellers and epidemiologists, and provides a basis for further work on pneumococcal dynamics, particularly to evaluate spatial dynamics and PCV strategies in the presence of important social heterogeneity and heterogeneity in health service (i.e. in this case vaccination) coverage, as outlined in subchapter 7.3.

In *Research paper 4* the combination of both carriage and contact data collected simultaneously in the same individuals provided a unique opportunity to better understand the

role of human interactions in the spread of respiratory infections, and *Streptococcus pneumoniae* in particular. Importantly, in the context of the work conducted for this thesis, *Research paper 4* was key to accurately parameterise the Nha Trang model, as it provided evidence supporting the use of physical contacts, rather than all contacts, in model parameterisation. Although it seems obvious that making a greater number of contacts might increase the risk of pneumococcal carriage – a hypothesis which also underlays most mathematical models, many public health interventions, and justifies why contact studies in various settings are important – the evidence to support this hypothesis is thin, consisting largely of ecological analyses or studies conducted over non-contemporary periods [27-32]. By shedding light on the type of contacts that are important for the transmission of *Streptococcus pneumoniae*, and the role of children in spreading pneumococci to older age groups, *Research paper 4* strengthens the scientific evidence for public health measures based upon the assumption that the frequency of close social contacts increases infection risk. It also provides crucial information to improve our understanding of transmission dynamics of pneumococci, and other pathogens transmitted through similar routes, and is central to adequately parameterise dynamic transmission models, including the model applied to the population of Nha Trang (*Research paper 5*).

Results from the first four papers all informed the mathematical model applied to the population of Nha Trang to explore the differential impact of PCV introduction strategies (*Research paper 5*). The choice of setting was based on considerations about (i) available data on both carriage, social contacts, and estimates of invasive disease, and (ii) the need for a model to address questions in a setting that has not yet introduced the vaccine – but is considering to do so –, and where results can therefore help guide policy decisions.

Research paper 5 provides the first dynamic model of pneumococcal transmission and control developed and parameterised for South-East Asia, where most countries are yet to introduce pneumococcal conjugate vaccination, and is therefore particularly relevant for policy decisions about PCV introduction in such settings. Results from the model showed a strong herd immunity effect of the vaccine, with near-complete elimination of serotypes targeted by the vaccine across all age groups within a decade of PCV introduction, in all vaccination strategies considered in the main scenario (i.e. 90% coverage in routine vaccination, and similarly in catch-up campaigns in <1 year olds, <2 year olds and <5 year olds). This predicted herd effect is in line with the trends seen in many settings following PCV introduction , including the experience in Europe and North America and South Africa, the Gambia [33-41] as well as following PCV10 introduction in Kenya [42].

The main results from the transmission model suggest that most of the benefit of catch-up campaigns occurs within the first 3 years of PCV introduction, with the highest impact at one year, when a catch-up strategy for <1 year olds would result in an 11% (95%CrI 9 – 14%) lower IPD incidence in children <5 years, compared to routine vaccination only, a strategy targeting children <2 years would reduce it by 25% (21 – 30 %), and a 38% (32 – 46%) reduction in incidence, compared to routine vaccination is predicted at one year with a catch-up strategy targeting all children under five years of age.

However, results of the model also show that catch-up campaign would only prevent more cases of IPD so long as they do not result in a vaccination introduction delayed by more than 6 months, for catch-up campaigns in <1 year olds, to about 18 months, for catch-up campaigns targeting all children under five years.

There are few available data on nasopharyngeal carriage post-PCV implementation in low- or middle-income settings. In Kilifi, Kenya, where PCV10 was introduced routinely in 2011

alongside a catch-up campaign in children under five years of age, VT carriage in <5y olds was reduced by about 64% 18 months after vaccine introduction, with a vaccination coverage for at least one dose of about 74% [42].   With such vaccination coverage – and assuming it is stable and reached on day 1 of the vaccine introduction –, the Nha Trang model predicts a decline in VT serotype carriage 18 months after vaccine introduction of 81% (95%CrI 65% - 92%). In a large PCV7 cluster randomized trial in the Gambia, the prevalence of VT carriage in 2 – 4y olds decreased by about 56% 22 months after vaccine introduction, with a two-dose catch-up campaign in <30m olds and a coverage of about 50%. In the Nha Trang model, such catch-up campaign with a coverage of 50% is predicted to reduce VT carriage by about 54% (95%CrI 41 – 64%) 22 months after PCV13 introduction.

However, it is notable that in a setting such as Nha Trang, where the sociodemographic and epidemiological characteristics differ from many low-income settings such as those mentioned above, VT carriage elimination and herd effects are predicted to be established relatively rapidly, even without a catch-up campaign. Although catch-up campaigns would decrease disease burden more rapidly, policy decisions about whether or not to conduct campaigns in such setting will crucially depend on results of economic analyses, to which the current model outputs could contribute. In its simplest way, a cost comparison between scenarios based on the number of additional vaccine doses required in catch-up campaigns compared to the additional impact of such campaigns on disease incidence, would provide initial answers. More comprehensive economic analyses would need to consider additional delivery, programme implementation, storage and other costs. For full cost-utility analyses, estimations of the quality adjusted life years (QALYs) gained for different scenarios and health outcomes considered are warranted. While such undertaking was beyond the scope of this thesis, it is worth noting that very few economic analyses of PCV programmes have so far been based on dynamic transmission models, and are based instead on static models where direct and indirect effects

of the model are fixed, or indirect effects not taken into account [43]. Even though medium and long term impacts can be estimated through simple static models [6], dynamic transmission models such as the model developed in *Research paper 5* are essential to appraise the population impact of different strategies when both direct and indirect effects are not stable, in the first few years post PCV introduction. Estimates from the Nha Trang model or similar models in other settings are important to inform economic analyses used to guide policy decisions about whether or not to include PCV catch-up campaigns.

7.2. Limitations

While the Nha Trang model, and the policy questions that were addressed through it, was a main objective of this thesis, there are a number of limitations to this model, including the available data used for its fitting and several assumptions inherent to its structure, some of which were already alluded to in the discussion of *Research paper 5*, and will therefore not be repeated in detail. Similarly, limitations to Research papers 1 to 4 were discussed in details in the respective papers, and will not be repeated here; rather, additional limitations of those studies in the specific context of their input into the Nha Trang analysis will be briefly discussed.

As already mentioned in the discussion of *Research paper 5*, the Nha Trang model was fitted to pre-vaccination data only, and predicted trends in post-vaccination depend on the efficacy of vaccination against carriage, estimated in *Research paper 2,* and the strength of competition, which determines the magnitude of replacement in carriage post vaccination. The evidence of strong competition between pneumococcal serotypes has been supported by observational studies finding little to no reduction in overall carriage rates after the introduction of routine vaccination [33-38], as well as evidence from mathematical models [44-48], mouse models [49]

and in vitro studies [50]. The competition parameter in the model was chosen to be similar to estimates from the UK and Finland [45, 51], which suggest very strong competition between VT and NVT and hence pronounced replacement effects after vaccination and the concomitant disappearance of VT from the nasopharynx.

The effect of replacement in the nasopharynx on disease rates is crucially dependent on the pathogenicity of the replacing serotypes. Although the possible impact on IPD was based on invasiveness data from England and Wales, it is impossible to predict the exact serotype that will fill the ecological niche created by vaccination, which may result in different degrees of pathogenicity. In the US the post PCV7 experience has been marked by a disproportional increase in serotype 19A, which is included in PCV13 [34]. In other countries, other serotypes such as serotypes 1 and serotypes 5, also included in PCV13, have disproportionally contributed to the post-vaccination IPD rates. Post-vaccination surveillance – or at least post-introduction monitoring of carriage prevalence and serotype distribution [6] – is therefore warranted to better predict long term post-vaccination trends.

Several simplifying assumptions were also made in the model structure. Although individuals may develop some immunity as a result of carriage [2, 52], the Nha Trang model assumed no natural immunity to pneumococcal infection. Some degree of immunity was nonetheless taken into account indirectly in the model, through decreasing the duration of carriage with age and differences in susceptibility to infection given contact. The model structure also overlooked the complexity of pneumococcal transmission, by assuming homogeneous characteristics for the group of VT and NVT serotypes, despite serotypes within each group differing in their pathogenicity, fitness and transmissibility [47, 53]. Serotype-specific post-PCV dynamics could therefore not be predicted, and it is likely that the differential impact of campaigns on some particular serotypes may differ, depending on the serotype-specific characteristics, including

255

prevalence and pathogenicity, and vaccine-related parameters such as efficacy and duration of protection. Exploring serotype-specific transmission parameters through modelling would help improve predictions, however there is currently no study that has done so.

Furthermore, there were several limitations in the parameters used in the Nha Trang model. Given that the model was fitted to pre-vaccination data only, post-vaccination trends were based on a range of assumptions about parameter values which are prone to uncertainty. Such uncertainty was taken into account to a large extent by assigning distributions to most parameters based on the best known estimates from the literature [20], and by exploring the sensitivity of the results to changes in some of the parameter values. Finally, in the Nha Trang model no heterogeneity was taken into account such as heterogeneity in vaccination coverage, individual susceptibility to disease, or population movements. Further work exploring the impact of population movement, particularly in the presence of heterogeneous vaccination coverage geographically [54], and differences in individual risk factors to disease (such as malnutrition or HIV), should further be explored, and is discussed in 7.3 (Ongoing research and areas for further work).

In *Research paper 2*, estimates of vaccine efficacy and duration of protection were almost entirely based on data from intervention studies based on PCV7 [20], whereas the transmission model was based on PCV13. Although immunological data suggest limited differences in efficacy estimates for the three or six additional serotypes in PCV10 and PCV13, as discussed earlier, whether aggregate efficacy estimates for PCV13 exactly match estimates obtained in *Research paper 2* [20] remains to be determined.

Finally, vaccine efficacy estimates for 'partial protection' in the Nha Trang model was based on a relatively small meta-analysis of incomplete (2 dose) vs complete routine schedules. However,

there are a number of efficacy estimates that are lacking for other vaccine dosages, and in particular single dose catch-up campaigns to expanded age groups. A better understanding of the efficacy and duration of protection of single dose schedules in toddlers is important for settings where a single-dose catch-up or single dose mass vaccination campaigns may be the most practical and pragmatic vaccination options.

7.3.   Ongoing work and areas for further research

Work undertaken as part of this thesis highlights a number of important additional questions on the epidemiology and transmission dynamics of pneumococci, and on how to best optimise and evaluate pneumococcal vaccination strategies. Work from this thesis also lays the basis for further projects, some of which are already ongoing. These include, among others:

1. A better understanding of the global epidemiology of *Streptococcus pneumoniae* carriage before the introduction of conjugate vaccines, for different geographical and sociodemographic settings, and by individual factors such as age and underlying health conditions.

2. Improving our understanding of the serotype-specific pathogenicity for invasive disease, but also non-invasive disease such as pneumonia and acute otitis media, the impact on which has so far been overlooked in most modelling studies, including the Nha Trang model.

3. Explore additional questions around the best use of PCV in high burden settings, including

    a. Humanitarian emergencies, where strategies need to consider reduced dose schedules, mass vaccination and expanded target age groups to account for the epidemiological conditions and reduced access to health services

    b. Settings where spatial coverage of PCV is highly heterogeneous

7.3.1. A Review of Streptococcus pneumoniae in nasopharyngeal carriage globally (RESPICAR study)

A better understanding of the global distribution of *S.pneumoniae* in carriage before PCV introduction would enable to better tailor, predict and monitor PCV vaccination strategies [9]. This is particularly relevant in low income settings where robust surveillance of pneumococcal disease remains a challenge. In addition, better estimates of the global distribution and prevalence of nasopharyngeal serotype carriage, and differences between settings, is important in the context of the development of new pneumococcal vaccines [55].

Adegbola et al. [15] undertook a systematic review of pneumococcal carriage in healthy children under five in low- and lower-middle income countries identified based on articles published in PubMed between 1990 and 2012. Their study focused primarily on carriage prevalence, rather than serotype distribution, and results suggest that the pooled carriage prevalence in low income settings tends to be higher than in lower-middle income countries. This may be due to the higher frequency of physical social contacts in such settings, as suggested by the findings from *Research paper 4*. And while Adegbola et al. [15] did not explore results for older age groups and adults, we know from the results of *Research paper 1* [1] that the prevalence of carriage in older children and adults is likely to be high in settings where the prevalence of carriage is high in children under five. In addition, as the results from *Research paper 1* suggests, NVT serotypes tend to be proportionally more prevalent in NP carriage as age increases [1], and under the assumption that this is related to serotype-specific acquired immunity to the most prevalent circulating serotypes, as mentioned earlier, it is expected that in high burden settings the shift towards a proportional increase in NVT carriage may happen earlier in life given the higher frequency of acquisition events. However, global and regional estimates of the serotype distribution of nasopharyngeal carriage have not been

published, and although results from *Research paper 2* [20] suggest a relatively constant distribution of vaccine-type serotypes in young children in a few settings, limited evidence exist about the serotype distribution in other settings.

Hence, we designed and implemented a large systematic review of *Streptococcus pneumoniae* carriage (RESPICAR) globally, in order to compile estimates of both the serotype distribution and carriage prevalence of pneumococci carried in the nasopharynx, across different age bands and geographic settings before PCV introduction (PI: Olivier le Polain, Research team: W John Edmunds (LSHTM), Stefan Flasche (LSHTM), Riya Moodley (LSHTM), Tim Pollington (LSHTM), Maria Deloria-Knoll (Johns Hopkins School of Public Health) and Kate O'Brien (Johns Hopkins School of Public Health)).

7.3.1.1. Description of the study methodology and preliminary results

The study, which was funded by the World Health Organisation, (WHO) started in late 2014 and is currently (November 2016) being finalised. Articles were retrieved from ten different databases, using a very inclusive search strategy and no restrictions by language, age group, time or region. The presence of underlying conditions (such as HIV or sickle cell disease) was not an exclusion criterion, but was recorded as part of the data entry, such as to conduct further analyses among those specific groups. Studies were included if they provided estimates from the community, and studies conducted among specific groups of patients with any form of pneumococcal-like illness (e.g. acute respiratory infection, acute otitis media, pneumonia) were recorded but not included for data extraction.

The first phase of the screening process, which was based on the screening of title and/or abstracts only, identified 14,584 studies (after deduplication of the initial 27,786 citations retrieved from the initial search). After a second screening phase based on full text, a total of

244 studies were identified which provided data on both carriage prevalence and serotype distribution among community members before PCV introduction.

Of the studies included for data extraction, 43 (18%) were from Africa, 69 (28%) from Asia, 76 (31%) from Europe, 33 (13%) from Latin America, 18 (7%) from North America and 5 (2%) from Oceania. These studies comprised a total of 511 different surveys or data collection phases, including 352 (69%) in children <5 years, 109 (21%) among 5-17 year olds and 50 (10%) among adults. We compiled information from a total of 70,901 isolates, including 51,299 isolates from children aged less than five years, 16,383 among children aged 5 – 17 years and 3,219 isolates from adults aged 18 years and (Figure 24).

Figure 24: Number of nasopharyngeal carriers with information on serotyping, by count
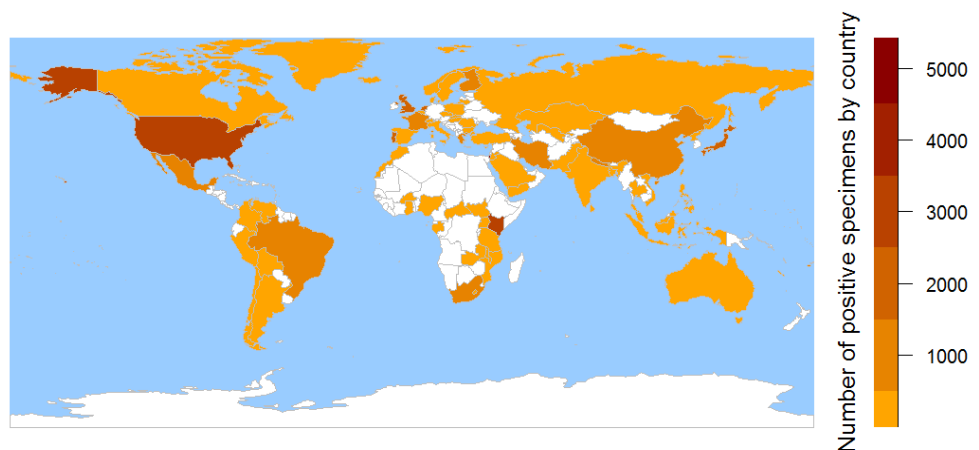
The analysis was based on separate models for serogroups and serotypes within serogroups (e.g. serogroup 19 includes serotypes 19A, 19B, 19C and 19F), whereby serogroups were modelled as a single multinomial distribution, and the distribution of serotypes within each serogroup was also independently modelled assuming a multinomial distribution. The analysis was performed in a Bayesian framework, to facilitate the inclusion of small or zero values for rarer serotypes, as well as provide a more natural weighting of the contribution of each study

than in a frequentist approach. The pooling of studies by age group and by geographic region was based upon (and interpreted as a function of) an analysis of heterogeneity. The latter included an assessment of the distribution of non-overlapping serotype- and/or serogroup-specific proportions between studies or groups of studies to compare, using their respective posterior samples.

The first main output of this project will be a very detailed and comprehensive description of the serotype distribution in carriage globally, by age group and geographical area. Preliminary results show differences in the serotype distribution and rank order of serotypes between different regions, but only moderate overall heterogeneity between countries, regions and continents.

Pooled data at the continent level show that among children<5 years of age the proportion of serotypes covered by PCV10 and PCV13 ranges from 47% (Asia) to 57% (Europe) for PCV10, and from 62% (Asia) to 74% (Europe) for PCV13. As hypothesised in *Research paper 1* [1], preliminary results also suggest a steady shift in the proportional distribution of serotypes from vaccine type serotypes in early childhood to non-vaccine type serotypes as age increases. Interestingly, while differences in serotype distributions are seen between settings, preliminary results suggest that a set of main prevalent serotypes are consistently found across settings. Importantly, this large review sets out the basis for further work, including subgroup analyses of the serotype distribution, work to explore the pathogenicity of specific strains, as well as work to further explore and predict the 'steady state' (i.e. medium to long term) impact of pneumococcal vaccine introduction in various settings using available and validated modelling and statistical tools [6, 56].

7.3.2. Serotype-specific pathogenicity for invasive and non-invasive disease

7.3.2.1. *Invasiveness of pneumococcal serotypes.*

The invasiveness is the probability of invasive disease per nasopharyngeal acquisition, which should be calculated as the incidence of disease divided by that of carriage, even though it is generally approximated by the incidence of disease divided by the prevalence of carriage (i.e. 'case:carrier ratios' as in *Research paper 5*). Invasiveness is known to vary widely between serotypes [57-59], and as such, the relative invasiveness of serotypes carried by the vaccine, compared to replacing serotypes, is one of the major factors that determine the success of a PCV programme, and that of a given vaccine formulation [60, 61]. Although invasiveness is thought to be a stable serotype-specific property in children under five [59, 62] – a hypothesis which also underpins the impact estimates of the Nha Trang model (*Research paper 5*) – there is only limited evidence to support this. Trotter et al. [58] found similar invasiveness patterns by serotypes in children <5 years and older children and adults in the UK, and such findings were recently corroborated by Weinberger et al. [57] based on data from Navajo communities in the US. However, it is likely that invasiveness differs between settings, due to differences in host factors, such as immunity patterns, and underlying susceptibility factors such as malnutrition or the prevalence of immunocompromising conditions such as HIV [63].

A better understanding of the stability – or not – of invasive serotypes between settings would therefore enable us to improve our understanding, and our predictions, of the impact of PCV globally, particularly for settings for which only carriage data are available [6], in addition to better forecasting the potential impact of new vaccine formulations (such as PCV15 [64]).

Data from RESPICAR, combined with data from existing systematic reviews of serotypes found in invasive disease globally [65], would enable to address some of those questions. This project is currently being taken forward (November 2016).

As mentioned earlier, one of the main limitations of the Nha Trang model (and most if not all modelling studies to date), is the lack of quantification of the potential impact of PCV vaccination schedules on non-bacteraemic pneumonia, despite it being the main cause of pneumococcal-related morbidity and mortality globally [66]. This is largely due to the lack of robust estimates about the propensity of serotypes (or at the minimum, the group of VT and NVT serotypes) to cause non-bacteraemic pneumonia, given that no easily available diagnostic test is currently available. The aetiology of pneumonia remains particularly challenging, as aetiological diagnosis generally requires invasive techniques, which are risk prone and may not alter the clinical management. This hampers the aetiological ascertainment of pneumonia cases, and in particular an understanding of the pathogenicity of different strains of pneumococci. A meta-analysis of four large randomised controlled trials of PCV in various settings (including Navajo communities in the US [67], The Gambia [68], Philippines [69] and South Africa [70]) estimated the efficacy of PCV against clinical pneumonia at 8%, and against radiological confirmed pneumonia of 36% [66]. These trials were not designed to determine if changes in pneumonia incidence would occur over time as a result of serotype replacement and may therefore overestimate the long-term effect of vaccination against this endpoint.

Studies using administrative databases based on robust designs, adjusting for trends or using control diagnoses, have been performed in Nicaragua [71], Brazil [72] and in Israel [73]. They all found substantial reductions in the burden of pneumonia hospitalisations [71, 72] or of community acquired alveolar pneumonia [73], albeit of varying magnitude. Indirect effects of PCV against pneumonia have also been observed [74]. Differences in magnitude may reflect the different PCV formulation in the studies, but most probably also result from differences between settings in the circulating pneumococci (which would be expected to change with time

since introduction of vaccination) and other pathogens causing clinical or radiological pneumonia.

Few studies however have explored the propensity of specific serotypes to cause pneumonia. The PERCH study (Pneumonia Etiology Research for Child Health) [75], which was conducted in seven different study sites (five African and two Asian countries) may shed more light on this in the near future. Using NP carriage as a proxy indicator of pneumonia aetiology has also been used in studies comparing serotype distribution in carriage among pneumonia patients to that of healthy individuals drawn from the same communities, with the assumption that carried serotypes among pneumonia patients are indicative of aetiology [76]. While this assumption is prone to many caveats, not least the fact that antibiotics provided to patients with pneumonia may reduce and/or change their nasopharyngeal flora, several studies seem to suggest that carried serotypes can be used as proxy measures of pneumonia aetiological agents, particularly if the density of carriage is high (with the assumption that a higher density would equate to a higher likelihood of causality) [76-79].

Thus, a compilation of studies with data on carriage prevalence, density, and serotype distribution in carriage among patients with pneumonia, compared to healthy individuals, may enable us to enhance our understanding of serotype-specific pneumonia pathogenicity, thereby also improving attempts to quantify the impact of PCV on the burden of pneumonia in various settings. Data from RESPICAR, together with additional data on pneumonia, would provide the opportunity to further this idea.

7.3.2.3. *Propensity of serotypes to cause acute otitis media.*

Finally, while acute otitis media (AOM) is a less severe pneumococcal disease manifestation, it is an important contributor to the burden of pneumococcal disease in children worldwide given its frequent occurrence, and even modest vaccine efficacy against AOM therefore results in

important reductions of the burden of the disease [80]. While aetiological diagnoses based on middle ear fluid (MEF) is generally specific, MEF is rarely collected as this procedure is seldom useful to guide clinical practice. In a recent model applied to data from Israel, Flasche et al. [7] have shown that monitoring serotypes in carriage can provide an alternative to both monitor and predict the impact of PCV on AOM. A further understanding of the AOM pathogenicity of various serotypes, based on either MEF serotyping or nasopharyngeal carriage among children with AOM, and NP carriage serotyping from the same community or individuals would enable, as for IPD and pneumonia, better predictions of the impact of PCV vaccination strategies on AOM.

7.3.3. Additional questions around the best use of PCV in high burden settings

7.3.3.1. *Applicability of the Nha Trang model results to other settings*

Although results from the Nha Trang model *(Research paper 5)* are informative for that particular settings, and perhaps other settings with similar epidemiological and socio-demographic characteristics, further work is warranted to explore the applicability of (or divergence from) the Nha Trang results when similar approaches are applied to other settings. This is particularly true for higher burden settings where carriage prevalence and social dynamics likely differ. Serotyping of the positive cultures from the nasopharyngeal carriage study in Sheema district (Nackers F, Cohuet S, le Polain de Waroux O, et al., (in preparation) and [5]) was only performed later, and modelling work applied to that setting was therefore not done in the context of this thesis. However, a model based on the same structure as the Nha Trang model was applied to both pre- and post-introduction NP carriage and IPD data from Kilifi, Kenya (Flasche et al., in preparation), another East African setting where both the prevalence of carriage, serotype distribution and socio-demographic characteristics are close to that of Sheema district, but differ substantially from that of Nha Trang [81]. Results from both

Kilifi and Nha Trang models were compared by running both model on the same parameters of vaccine efficacy and duration of protection to explore similarities and differences in the differential impact of catch-up campaigns in those two different epidemiological contexts, the results of which will further feed into economic analyses to help help Gavi, the Vaccine Alliance to evaluate whether pneumcoccal catch-up campaigns should be included in their investment strategy.

7.3.3.2. *Impact of PCV in the presence of spatially heterogeneous vaccination coverage*

Little attention has been given to the impact of spatial or social heterogeneity in vaccination coverage and transmission on the impact of PCV. There is good evidence from various low-income settings that vaccination coverage for routine childhood vaccines often decreases in areas located further away from health centres, as has been reported in studies from Bangladesh [82], Mozambique [83], Kenya [84], Tanzania [54] and Malawi [85]. The vaccination coverage needed to achieve herd immunity in populations where vaccine coverage varies spatially will depend on the transmission dynamics within different geographic areas and the degree of mixing between geographical areas [86]. Although such information is likely to be setting-specific, assessing the effect of spatial heterogeneity and between-area mixing on the transmission dynamics of *Streptococcus pneumoniae* would enable to better understand the population-wide impact of PCVs at different coverage levels and the possible advantage (or not) of PCV outreach activities in under-vaccinated areas under various scenarios of spatial heterogeneity in vaccination coverage. Even in areas where health facilities are reasonably easy to access, children living further away from such facilities tend to be inadequately vaccinated, and this was shown recently in the context of PCV introduction in Malawi [85]. The degree of 'spillover effect' from adequately vaccinated areas into poorly vaccinated areas in a rural African setting is not known, and such information would help better tailor supplementary

immunisation activities, if required, in areas of low vaccine uptake. Given the availability of data on both nasopharyngeal carriage (including serotype distribution) across all age groups, and social contact data with spatially explicit information (*Research paper 3*) from Sheema district, a meta-population model could be constructed, in which transmission parameters derived from the contact study could be used to calculate the fraction of social contacts occurring within defined geographical patches and between geographic patches.

Scenarios of vaccination coverage such as systematic differences in vaccination coverage across patches, with vaccination coverage in patches decreasing proportionally to the patch's distance to a health centre, as well as extremes of vaccination coverage, could further be explored.

7.3.3.3.   *Optimal PCV strategies in humanitarian emergencies.*

It is unlikely that vaccination schedules adapted to stable settings are adequate for crisis settings, where population displacement, overcrowding and intensified social mixing, the disruption of preventative and curative health services, as well as increased prevalence of malnutrition may impact disease dynamics, resulting in excess disease and death, altered age distributions and higher herd immunity thresholds [87-89]. In many crisis-affected settings routine immunisation services are heavily disrupted, with vaccination strategies relying on supplemental immunisation activities (SIAs) to increase coverage, also among age groups not usually targeted through routine immunisation programmes.

As the number of internally and externally displaced individuals due to complex emergencies is increasing and is higher than ever, the need for appropriate and optimised vaccination strategies is ever more critical. The WHO's Strategic Advisory Group of Experts on Immunization in Emergencies recommends for vaccination strategies to be adapted in such contexts, including options such as mass vaccination campaigns, reduced dose schedules and

vaccination to expanded age groups in order to promptly prevent excess death and mortality related to the emergency [90].

In 2013 refugees in Yida, one of the largest refugee camps in South Sudan, were some of the first refugees or internally displaced individuals to receive PCV [91]. In September 2016, GlaskoSmithKline (GSK) agreed to lower the price of PCV for crisis-affected populations served by humanitarian organisations, offering hope that PCV will now be considered more consistently in such settings.

However, PCV will remain expensive compared with other immunisation programmes and estimating its impact in stable settings is complex enough. Hence, many questions remain around the best use of PCV in complex emergencies, not least because crises settings all differ. Nevertheless, a number of important questions are applicable to many including (i) strategies for one-off or regular supplementary immunisation activities (SIAs) in areas where overall vaccine uptake is low, (ii) the impact of single- versus multiple-dose SIAs, and (iii) the impact of an influx of unvaccinated individuals (e.g. refugees or internally displaced) on transmission in (partially) vaccinated host communities. Work from this thesis will be pursued to address some of the questions highlighted above.

In Figure 25 for example, an SIA was assumed for all children <2 years or <5 years 18 months after the introduction of PCV, based on the Nha Trang model, assuming a very low routine vaccination of 20% but a higher (80%) coverage through SIA.

While the setting and model is not appropriate for any adequate inference to an emergency setting, it illustrates for a relatively low burden setting a one-off mass vaccination strategy in the context of very low coverage of the childhood immunisation programme may result in substantial reductions in VT carriage for a number of years.

Further research on the optimisation of PCV strategies in such settings is urgently warranted.

Figure 25: VT carriage following a SIA (coverage 80%) conducted 18 months after PCV introduction (routine coverage of 20%)

Age <5 yrs
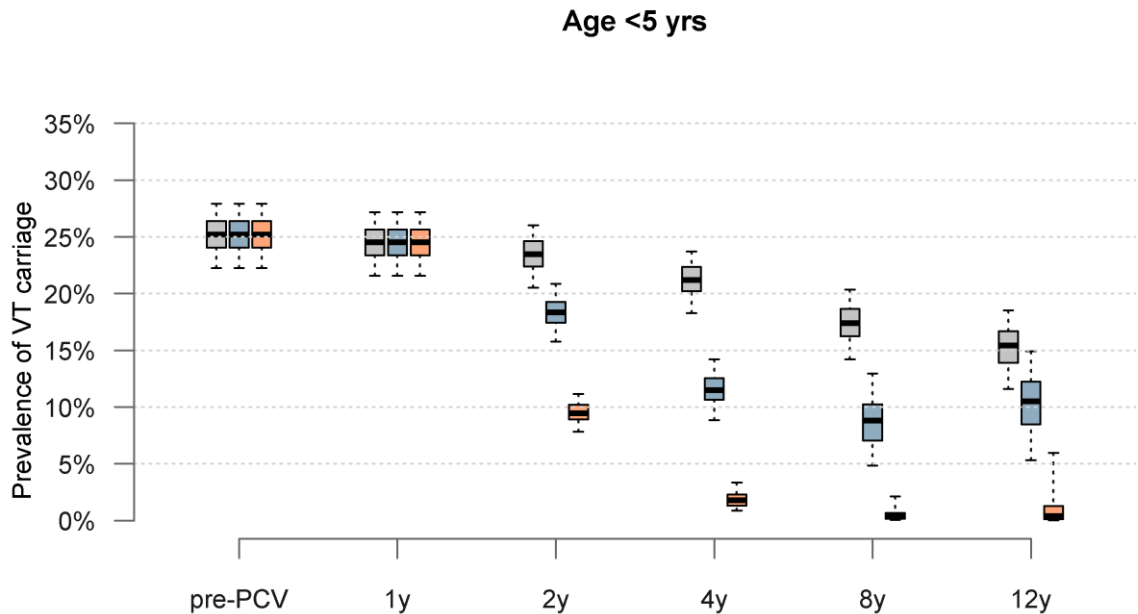
Legend: grey: RV (20% coverage), blue: SIA in <2years; red: SIA <5 years (SIA conducted at 18 months, 80% coverage)

7.4. Conclusion

The main focus of this thesis was to explore vaccination programme options at the introduction of the pneumococcal conjugate vaccine (PCV) in Vietnam through a dynamic model of disease transmission, and more specifically the potential impact on disease of catch-up campaigns targeting various age groups. Few countries – in particular low and lower-middle-income countries – have conducted such campaigns and there is therefore limited data to inform policy. Countries supported by Gavi, the Vaccine Alliance, have so far not been able to undertake catch-up campaigns at PCV introduction over concerns that they would lead to vaccine introduction delays, given supply and logistical constraints.

Work from this thesis provides the first dynamic model of pneumococcal disease applied to a population in South-East Asia, where most countries have not yet introduced PCV, and where the burden of pneumococcal disease is relatively high. The availability of data on human social mixing patterns and of nasopharyngeal carriage prevalence and serotype distribution of pneumococci in the same population allowed to thoroughly parameterise a model in that setting. Additional studies on the age-specific epidemiology of nasopharyngeal carriage, the vaccine efficacy against carriage and its waning, and social mixing patterns relevant to the spread of pneumococci directly contributed to improve the quality of the transmission model, but also provided important insight into the epidemiology and transmission dynamics of *S.pneumoniae*, which has implications for vaccination programmes.

Despite a number of limitations, results from this thesis feed directly into the current debates about vaccine introduction strategies, particularly in South-East Asia, offering insights into the impact of different PCV programme options, and providing estimates that could inform vaccine introduction strategies in such settings.

Work from this thesis also lays a basis on which to build and pursue further research around the epidemiology, transmission dynamics, and optimisation of PCV vaccination strategies, particularly in high burden settings where pneumococcal disease still remains one of the most important causes of morbidity and mortality.

References

[1] Le Polain de Waroux O, Flasche S, Prieto-Merino D, Edmunds WJ. Age-dependent prevalence of nasopharyngeal carriage of streptococcus pneumoniae before conjugate vaccine introduction: a prediction model based on a meta-analysis. PloS one. 2014;9:e86136.

[2] Weinberger DM, Dagan R, Givon-Lavi N, Regev-Yochay G, Malley R, Lipsitch M. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. The Journal of infectious diseases. 2008;197:1511-8.

[3] Cobey S, Lipsitch M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. Science. 2012;335:1376-80.

[4] Joloba ML, Bajaksouzian S, Palavecino E, Whalen C, Jacobs MR. High prevalence of carriage of antibiotic-resistant Streptococcus pneumoniae in children in Kampala Uganda. International journal of antimicrobial agents. 2001;17:395-400.

[5] Nanjebe D, Nyehangane D, Langendorf D, Mwanga-Amumpaire J, Cohuet S, Scott A, et al. A Baseline pneumococcal carriage survey prior to the introduction of PCV10 in Uganda. 10th International Symposium on Pneumococci and Pneumococcal Disease (ISPPD-10). Glasgow2016.

[6] Flasche S, Le Polain de Waroux O, O'Brien KL, Edmunds WJ. The serotype distribution among healthy carriers before vaccination is essential for predicting the impact of pneumococcal conjugate vaccine on invasive disease. PLoS computational biology. 2015;11:e1004173.

[7] Flasche S, Givon-Lavi N, Dagan R. Using Pneumococcal Carriage Data to Monitor Postvaccination Changes in the Incidence of Pneumococcal Otitis Media. American journal of epidemiology. 2016;184:652-9.

[8] Pneumococcal Carriage C, Kllyhty H, Auranen K, Dagan R, Goldblatt D, O'Brien K, et al. Case for carriage. Vaccine. 2013;32:180-203.

[9] Goldblatt D, Ramakrishnan M, O'Brien K. Using the impact of pneumococcal vaccines on nasopharyngeal carriage to aid licensing and vaccine implementation; a PneumoCarr meeting report March 27-28, 2012, Geneva. Vaccine. 2013;32:146-52.

[10] Davis S, Deloria-Knoll M, Kassa HT, O'Brien KL. Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: Review of evidence on indirect effects. Vaccine. 2013.

[11] van Werkhoven CH, Bonten MJ. The Community-Acquired Pneumonia immunization Trial in Adults (CAPiTA): what is the future of pneumococcal conjugate vaccination in elderly? Future Microbiol. 2015;10:1405-13.

[12] Chalmers JD, Campling J, Dicker A, Woodhead M, Madhava H. A systematic review of the burden of vaccine preventable pneumococcal disease in UK adults. BMC Pulm Med. 2016;16:77.

[13] Pilishvili T, Bennett NM. Pneumococcal Disease Prevention Among Adults: Strategies for the Use of Pneumococcal Vaccines. American journal of preventive medicine. 2015;49:S383-90.

[14] Flasche S, Hoek AJv, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoS medicine. 2011;8.

[15] Adegbola RA, DeAntonio R, Hill PC, Roca A, Usuf E, Hoet B, et al. Carriage of Streptococcus pneumoniae and Other Respiratory Bacterial Pathogens in Low and Lower-Middle Income Countries: A Systematic Review and Meta-Analysis. PloS one. 2014;9:e103293.

[16] Usuf E, Bottomley C, Adegbola RA, Hall A. Pneumococcal carriage in sub-Saharan Africa--a systematic review. PloS one. 2014;9:e85001.

[17] Central Office for Statistics. Age-sex structure and marital status of the population of Viet Nam. Hanoi: Ministry of Planning and Investment; 2011.

[18] Uganda Bureau of Statistics. Uganda Demographic and Health Survey Kampala2011.

[19] Gambia Bureau of Statistics. Demographic and Health Survey, The Gambia. Banjul, The Gambia2013.

[20] Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. The Pediatric infectious disease journal. 2015;34:858-64.

[21] Choi Y, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, et al. 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PloS one. 2011;28.

[22] Lucero MG, Dulalia VE, Nillos LT, Williams G, Parreno RA, Nohynek H, et al. Pneumococcal conjugate vaccines for preventing vaccine-type invasive pneumococcal disease and X-ray defined pneumonia in children less than two years of age. Cochrane database of systematic reviews. 2009:CD004977.

[23] Simell B, Auranen K, Kayhty H, Goldblatt D, Dagan R, O'Brien KL, et al. The fundamental link between pneumococcal carriage and disease. Expert review of vaccines. 2012;11:841-55.

[24] Vu HT, Yoshida LM, Suzuki M, Nguyen HA, Nguyen CD, Nguyen AT, et al. Association between nasopharyngeal load of Streptococcus pneumoniae, viral coinfection, and radiologically confirmed pneumonia in Vietnamese children. PediatrInfectDisJ2011. p. 11-8.

[25] Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoSMed. 2008;5:e74.

[26] Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and urban areas of southern China. Proceedings Biological sciences / The Royal Society. 2014;281:20140268.

[27] Santermans E, Goeyvaerts N, Melegaro A, Edmunds WJ, Faes C, Aerts M, et al. The social contact hypothesis under the assumption of endemic equilibrium: Elucidating the transmission potential of VZV in Europe. Epidemics. 2015;11:14-23.

[28] Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. American journal of epidemiology. 2006;164:936-44.

[29] Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ. What types of contacts are important for the spread of infections?: using contact survey data to explore European mixing patterns. Epidemics. 2011;3:143-51.

[30] Kwok KO, Cowling BJ, Wei VW, Wu KM, Read JM, Lessler J, et al. Social contacts and the locations in which they occur as risk factors for influenza infection. Proceedings Biological sciences / The Royal Society. 2014;281:20140709.

[31] Kucharski AJ, Kwok KO, Wei VW, Cowling BJ, Read JM, Lessler J, et al. The contribution of social behaviour to the transmission of influenza A in a human population. PLoS pathogens. 2014;10:e1004206.

[32] Rohani P, Zhong X, King AA. Contact network structure explains the changing epidemiology of pertussis. Science. 2010;330:982-5.

[33] Flasche S, van Hoek AJ, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoSMed. 2011;8:e1001017.

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

[35] Spijkerman J, van Gils EJ, Veenhoven RH, Hak E, Yzerman EP, van der Ende A, et al. Carriage of Streptococcus pneumoniae 3 years after start of vaccination program, the Netherlands. EmergInfectDis. 2011;17:584-91.

[36] Ghaffar F, Barton T, Lozano J, Muniz LS, Hicks P, Gan V, et al. Effect of the 7-valent pneumococcal conjugate vaccine on nasopharyngeal colonization by Streptococcus pneumoniae in the first 2 years of life. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2004;39:930-8.

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

[38] Hennessy TW, Singleton RJ, Bulkow LR, Bruden DL, Hurlburt DA, Parks D, et al. Impact of heptavalent pneumococcal conjugate vaccine on invasive disease, antimicrobial resistance and colonization in Alaska Natives: progress towards elimination of a health disparity. Vaccine. 2005;23:5464-73.

[39] Hammitt LL, Bruden DL, Butler JC, Baggett HC, Hurlburt DA, Reasonover A, et al. Indirect effect of conjugate vaccine on adult carriage of Streptococcus pneumoniae: an explanation of trends in invasive pneumococcal disease. JInfectDis. 2006;193:1487-94.

[40] Nzenze SA, Shiri T, Nunes MC, Klugman KP, Kahn K, Twine R, et al. Temporal changes in pneumococcal colonization in a rural African community with high HIV prevalence following routine infant pneumococcal immunization. The Pediatric infectious disease journal. 2013;32:1270-8.

[41] Roca A, Hill PC, Townend J, Egere U, Antonio M, Bojang A, et al. Effects of community-wide vaccination with PCV-7 on pneumococcal nasopharyngeal carriage in the Gambia: a cluster-randomized trial. PLoSMed. 2011;8:e1001107.

[42] Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. The Lancet Global health. 2014;2:e397-405.

[43] Saokaew S, Rayanakorn A, Wu DB, Chaiyakunapruk N. Cost Effectiveness of Pneumococcal Vaccination in Children in Low- and Middle-Income Countries: A Systematic Review. Pharmacoeconomics. 2016.

[44] Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. PloS one. 2012;7:e39927.

[45] Choi YH, Jit M, Gay N, Andrews N, Waight PA, Melegaro A, et al. 7-Valent pneumococcal conjugate vaccination in England and Wales: is it still beneficial despite high levels of serotype replacement? PLoSOne. 2011;6:e26190.

[46] Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, Gay NJ. Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. BMCInfectDis. 2010;10:90.

[47] Lipsitch M, Abdullahi O, D'Amour A, Xie W, Weinberger DM, Tchetgen Tchetgen E, et al. Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. Epidemiology. 2012;23:510-9.

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

[49] Lipsitch M, Dykes JK, Johnson SE, Ades EW, King J, Briles DE, et al. Competition among Streptococcus pneumoniae for intranasal colonization in a mouse model. Vaccine. 2000;18:2895-901.

[50] Rajam G, Jackson D, Pilishvili T, Whitney CG, Facklam RR, Carlone GM, et al. An in vitro model to assess pneumococcal adherence to nasopharyngeal cells under competition conditions. Journal of microbiological methods. 2007;70:219-26.

[51] Auranen K, Mehtala J, Tanskanen A, Kaltoft S. Between-strain competition in acquisition and clearance of pneumococcal carriage--epidemiologic evidence from a longitudinal study of day-care children. AmJEpidemiol. 2010;171:169-76.

[52] Pomat WS, Lehmann D, Sanders RC, Lewis DJ, Wilson J, Rogers S, et al. Immunoglobulin G antibody responses to polyvalent pneumococcal vaccine in children in the highlands of Papua New Guinea. Infection and immunity. 1994;62:1848-53.

[53] Hogberg L, Geli P, Ringberg H, Melander E, Lipsitch M, Ekdahl K. Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. Journal of clinical microbiology. 2007;45:948-52.

[54] Le Polain de Waroux O, Schellenberg JR, Manzi F, Mrisho M, Shirima K, Mshinda H, et al. Timeliness and completeness of vaccination and risk factors for low and late vaccine uptake in young children living in rural southern Tanzania. International health. 2013;5:139-47.

[55] Moffitt KL, Malley R. Next generation pneumococcal vaccines. Current opinion in immunology. 2011;23:407-13.

[56] Weinberger DM, Harboe ZB, Flasche S, Scott JA, Lipsitch M. Prediction of serotypes causing invasive pneumococcal disease in unvaccinated and vaccinated populations. Epidemiology. 2011;22:199-207.

[57] Weinberger D, Grant LR, Weatherholtz R, Warren JL, O'Brien K, Hammitt L. Relating Pneumococcal Carriage Among Children to Disease Rates Among Adults Before and After the Introduction of Conjugate Vaccines. American journal of epidemiology. 2016;183 1055-62.

[58] Trotter CL, Waight P, Andrews NJ, Slack M, Efstratiou A, George R, et al. Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. The Journal of infection. 2010;60:200-8.

[59] Brueggemann AB, Spratt BG. Geographic distribution and clonal diversity of Streptococcus pneumoniae serotype 1 isolates. Journal of clinical microbiology. 2003;41:4966-70.

[60] Nurhonen M, Auranen K. Optimal serotype compositions for Pneumococcal conjugate vaccination under serotype replacement. PLoS computational biology. 2014;10:e1003477.

[61] Yildirim I, Hanage WP, Lipsitch M, Shea KM, Stevenson A, Finkelstein J, et al. Serotype specific invasive capacity and persistent reduction in invasive pneumococcal disease. Vaccine. 2010;29:283-8.

[62] Brueggemann AB, Griffiths DT, Meats E, Peto T, Crook DW, Spratt BG. Clonal relationships between invasive and carriage Streptococcus pneumoniae and serotype- and clone-specific differences in invasive disease potential. The Journal of infectious diseases. 2003;187:1424-32.

[63] Backhaus E, Berg S, Andersson R, Ockborn G, Malmstrom P, Dahl M, et al. Epidemiology of invasive pneumococcal infections: manifestations, incidence and case fatality rate correlated to age, gender and risk factors. BMC infectious diseases. 2016;16:367.

[64] McFetridge R, Meulen AS, Folkerth SD, Hoekstra JA, Dallas M, Hoover PA, et al. Safety, tolerability, and immunogenicity of 15-valent pneumococcal conjugate vaccine in healthy adults. Vaccine. 2015;33:2793-9.

[65] Johnson HL, Deloria-Knoll M, Levine OS, Stoszek SK, Freimanis Hance L, Reithinger R, et al. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. PLoS medicine. 2010;7.

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

[67] O'Brien KL, Moulton LH, Reid R, Weatherholtz R, Oski J, Brown L, et al. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. Lancet. 2003;362:355-61.

[68] Cutts FT, Zaman SM, Enwere G, Jaffar S, Levine OS, Okoko JB, et al. Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. Lancet. 2005;365:1139-46.

[69] Lucero MG, Nohynek H, Williams G, Tallo V, Simoes EA, Lupisan S, et al. Efficacy of an 11-valent pneumococcal conjugate vaccine against radiologically confirmed pneumonia among children less than 2 years of age in the Philippines: a randomized, double-blind, placebo-controlled trial. The Pediatric infectious disease journal. 2009;28:455-62.

[70] Klugman KP, Madhi SA, Huebner RE, Kohberger R, Mbelle N, Pierce N. A trial of a 9-valent pneumococcal conjugate vaccine in children with and those without HIV infection. The New England journal of medicine. 2003;349:1341-8.

[71] Becker-Dreps S, Amaya E, Liu L, Moreno G, Rocha J, Briceno R, et al. Changes in childhood pneumonia and infant mortality rates following introduction of the 13-valent pneumococcal conjugate vaccine in Nicaragua. The Pediatric infectious disease journal. 2014;33:637-42.

[72] Sgambatti S, Minamisava R, Afonso ET, Toscano CM, Bierrenbach AL, Andrade AL. Appropriateness of administrative data for vaccine impact evaluation: the case of pneumonia hospitalizations and pneumococcal vaccine in Brazil. Epidemiology and infection. 2015;143:334-42.

[73] Givon N, Dagan R, Ben-Shimol S, Segal N, Bar-Ziv J, Greenberg D. Community Acquired Alveolar Pneumonia (CAAP) Incidence Rate Dynamics in Two Distinct Populations in Southern Israel Following PCV7/PCV13 sequential introduction. 10th International Symposium on Pneumococci and Pneumococcal Disease (ISPPD-10). Glasgow, UK2016.

[74] Loo JD, Conklin L, Fleming-Dutra KE, Deloria Knoll M, Park DE, Kirk J, et al. Systematic review of the effect of pneumococcal conjugate vaccine dosing schedules on prevention of pneumonia. The Pediatric infectious disease journal. 2014;33 Suppl 2:S140-51.

[75] Levine OS, Bhat N, Crawley J, Deloria-Knoll M, DeLuca AN, Driscoll AJ, et al. Pneumonia etiology research for child health. Introduction. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2012;54 Suppl 2:S87-8.

[76] Greenberg D, Givon-Lavi N, Newman N, Bar-Ziv J, Dagan R. Nasopharyngeal carriage of individual Streptococcus pneumoniae serotypes during pediatric pneumonia as a means to estimate serotype disease potential. Pediatric Infectious Disease Journal. 2011;30:227-33.

[77] Chappuy H, Keitel K, Gehri M, Tabin R, Robitaille L, Raymond F, et al. Nasopharyngeal carriage of individual Streptococcus pneumoniae serotypes during pediatric radiologically confirmed community acquired pneumonia following PCV7 introduction in Switzerland. BMC infectious diseases. 2013;13.

[78] Albrich WC, Madhi SA, Adrian PV, van Niekerk N, Mareletsi T, Cutland C, et al. Use of a rapid test of pneumococcal colonization density to diagnose pneumococcal pneumonia. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2012;54:601-9.

[79] Albrich WC, Madhi SA, Adrian PV, van Niekerk N, Telles JN, Ebrahim N, et al. Pneumococcal colonisation density: a new marker for disease severity in HIV-infected adults with pneumonia. BMJ Open. 2014;4:e005953.

[80] Fitzwater SP, Chandran A, Santosham M, Johnson HL. The worldwide impact of the seven-valent pneumococcal conjugate vaccine. The Pediatric infectious disease journal. 2012;31:501-8.

[81] Scott JA, Bauni E, Moisi JC, Ojal J, Gatakaa H, Nyundo C, et al. Profile: The Kilifi Health and Demographic Surveillance System (KHDSS). International journal of epidemiology. 2012;41:650-7.

[82] Bhuiya A, Bhuiya I, Chowdhury M. Factors affecting acceptance of immunization among children in rural Bangladesh. Health policy and planning. 1995;10:304-12.

[83] Cutts FT, Rodrigues LC, Colombo S, Bennett S. Evaluation of factors influencing vaccine uptake in Mozambique. International journal of epidemiology. 1989;18:427-33.

[84] Ndiritu M, Cowgill KD, Ismail A, Chiphatsi S, Kamau T, Fegan G, et al. Immunization coverage and risk factors for failure to immunize within the Expanded Programme on Immunization in Kenya after introduction of new Haemophilus influenzae type b and hepatitis b virus antigens. BMC public health. 2006;6:132.

[85] Mvula H, Heinsbroek E, Chihana M, Crampin AC, Kabuluzi S, Chirwa G, et al. Predictors of Uptake and Timeliness of Newly Introduced Pneumococcal and Rotavirus Vaccines, and of Measles Vaccine in Rural Malawi: A Population Cohort Study. PloS one. 2016;11:e0154997.

[86] Grassly NC, Fraser C. Mathematical models of infectious disease transmission. Nature reviews Microbiology. 2008;6:477-87.

[87] Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. Conflict and health. 2010;4:3.

[88] Checchi FG, M,; Grais, R.; Mills, E. . Public Health in crisis-affected populations: a practical guide for decision makers.   . Humanitarian Policy Network London: Overseas Development Institute; 2007.

[89] Spiegel PB, Checchi F, Colombo S, Paik E. Health-care needs of people affected by conflict: future trends and changing frameworks. Lancet. 2010;375:341-5.

[90] (WHO) WHO. Vaccination in acute humanitarian emergencies: a framework for decision-making. Geneva: WHO; 2013.

[91] Moszynski P. Refugees in South Sudan to receive pneumococcal vaccine after delays over price. Bmj. 2013;347:f5042.