Le Polain de Waroux, O; Flasche, S; Prieto-Merino, D; Edmunds, WJ (2014) Age-Dependent Prevalence of Nasopharyngeal Carriage of Streptococcus pneumoniae before Conjugate Vaccine Introduction: A Prediction Model Based on a Meta-Analysis. PloS one, 9 (1). e86136. ISSN 1932-6203 DOI: https://doi.org/10.1371/journal.pone.0086136

Downloaded from: http://researchonline.lshtm.ac.uk/1496173/

DOI: 10.1371/journal.pone.0086136

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
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 Waroux¹*, Stefan Flasche¹, David Prieto-Merino², W. John Edmunds¹

¹ Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom, ²Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom

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 <5 y 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 children (<5 y) and both that in school aged children (5–17 y 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 <5 y among 5–17 y olds and adults respectively. We provide a prediction model to infer the carriage prevalence and serotype distribution in 5–17 y 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.


Editor: Hiroshi Nishiura, The University of Tokyo, Japan

Received September 30, 2013; Accepted December 4, 2013; Published January 23, 2014

Copyright: © 2014 Le Polain de Waroux et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Olivier Le Polain de Waroux was supported by a doctoral research fellowship from the AXA Research fund for this piece of work. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: olivier.lepolain@lshtm.ac.uk

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 <5 y and that in older children (3–17 y olds) and adults (≥18 y 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.

Methods

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 Checklist S1) and MOOSE guidelines [13], which compile guidelines for the reporting of meta-analysis of observational studies.

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.

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 File S1.

Analysis

We considered the following age groups: <1 y (infants); <5 y (pre-school children including infants); 5–17 y (school aged children) and ≥18 y (adults). Because age groups were not standardised between studies, the category of <5 y olds included studies reporting estimates in <4 y olds as well as studies reporting estimates in <6 y 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].

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

\[
\text{P}_{\text{Ai}} = \beta_0 + \beta_1 \times \text{P}_{\text{Ci}} + \epsilon_i,
\]

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

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

We provided prior distributions for the parameters of interest $\beta_0$, $\beta_1$, and $\sigma$ and assigned uniform uninformative priors to $\beta_0$ (unif $(-1,1)$) and $\beta_1$ (unif $(-5,5)$) and to $\sigma$ (unif $[0,4]$).

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 $<5$ y of age and the % $<15$ y 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.

**Age-dependent distribution of VT serotypes.** We then explored (i) the distribution of the group of VT serotypes among carried serotypes in children $<5$ y 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 $5-17$ y 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 $\sigma$ 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 File S2 and is fully annotated. In addition, File S2 also provides an opportunity for readers to obtain posterior distributions of the carriage prevalence in $5-17$ y olds and adults based on study-specific estimates of nasopharyngeal carriage in $<5$ y olds, making it possible to use this as a carriage prediction tool based on specific data of carriage in $<5$ y olds.

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

**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 $<5$ y, 3936 school aged children and 8324 adults.

**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, including details about how individuals were selected in each age group, 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 $5-17$ y olds as a function of that in either $<5$ y olds or $<1$ y olds. 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. In all analyses and for all models considered the values of $\beta_1$ coefficients were unaffected and the coefficients of the variable explored were centred around zero.

We also explored whether prediction estimates for the adult population differed with studies ($n = 11$) where a random community or hospital sample of adults was taken compared to studies ($n = 8$) based on convenience sampling of parents, guardians or day care centre staff looking after children included (see Table 1 for further details). We found no difference in the prediction estimates obtained in subgroup analyses. Figure S1 shows the plot for those analyses.

Specific estimates in older adults were provided in two studies included in the analysis, one in Israel with data in $>65$ y olds [16] and another in Kenya with data in $>50$ y 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 18–65 y vs. 4.6% in $>65$ y, p = 0.573 and Kenya: 5.6% in 20–49 y olds vs. 4.6% in $\geq$50 y olds, p = 0.719).

**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. In all studies included serotyping was performed using standard methods such as latex agglutination and capsular swelling (Quellung reaction).

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 $<2$ y olds rather than $<5$ y 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.
A total of 8,886 citations were identified:
- 4,755 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

Figure 1. Flow chart of the study selection process.
doi:10.1371/journal.pone.0086136.g001
<table>
<thead>
<tr>
<th>Authors</th>
<th>Country or region</th>
<th>Study population characteristics by age group</th>
<th>No carriers total individuals (%)</th>
<th>by age group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdullahi et al. (2008) [17]</td>
<td>Kenya</td>
<td>C</td>
<td>58/98 (59%) 1 y olds; 198/349 (57%) 5 y olds; 55/213 (55%) 5–17 y olds; 16/302 (5%) 18 y</td>
<td>&lt;1 y &lt;2 y 5–17 y &gt;18 y</td>
</tr>
<tr>
<td>Adetifa IM., et al. (2012) [34]</td>
<td>Nigeria</td>
<td>C</td>
<td>143/193 (74%) 1 y olds; 375/524 (72%) 5 y olds; 63/125 (53%) 5–14 y olds; 90/356 (25%) 15–19 y olds; 7/841 (11%) 20 y</td>
<td>&lt;1 y &lt;5 y 5–14 y &gt;15 y</td>
</tr>
<tr>
<td>Bello Gonzales et al. (2010) [35]</td>
<td>Venezuela</td>
<td>C</td>
<td>58/84 (69%) 1 y olds; 7/64 (11%) 5 y olds; 18/174 (10%) 18 y</td>
<td>&lt;1 y &lt;5 y 18 y</td>
</tr>
<tr>
<td>Cekmez et al. (2009) [36]</td>
<td>Turkey</td>
<td>H</td>
<td>0/125 (0%)* 4 y olds; 25/375 (7%) 4–12 y olds; 61/1471 (5%) 19 y</td>
<td>&lt;1 y &lt;5 y 19 y</td>
</tr>
<tr>
<td>Chen et al. (2007) [37]</td>
<td>Taiwan</td>
<td>C&amp;H</td>
<td>25/94 (27%) 3–4 y olds; 18/196 (9%) 5–17 y olds; 3/461 (1%) 18 y</td>
<td>&lt;1 y &lt;3 y 3–4 y 5–17 y 18 y</td>
</tr>
<tr>
<td>Dagan et al. (2000) [38]</td>
<td>Israel</td>
<td>C</td>
<td>59/84 (70%) 6 y olds; 71/199 (35%) 6–15 y olds; 27/174 (15%) 16 y</td>
<td>&lt;1 y &lt;5 y 6–15 y 16 y</td>
</tr>
<tr>
<td>Darboe MK., et al. (2007) [39]</td>
<td>Gambia</td>
<td>C</td>
<td>143/196 (73%) 1 y olds; 26/196 (13%) 18 y</td>
<td>&lt;1 y &lt;3 y 18 y</td>
</tr>
<tr>
<td>Dhakal et al. (2010) [40]</td>
<td>India</td>
<td>H</td>
<td>18/79 (23%) 4 y olds; 26/120 (22%) 4–12 y olds; 12/304 (4%) 18 y</td>
<td>&lt;1 y &lt;5 y 18 y</td>
</tr>
<tr>
<td>Granat et al. (2007) [41]</td>
<td>Bangladesh</td>
<td>C</td>
<td>49/99 (49%) 1 y olds; 86/172 (50%) 5 y olds; 42/117 (36%) 5–18 y olds; 86/172 (50%) 19 y</td>
<td>&lt;1 y &lt;5 y 5–18 y 19 y</td>
</tr>
<tr>
<td>Greenberg et al. (2004) [42]</td>
<td>Israel</td>
<td>H</td>
<td>147/216 (68%) 5 y olds; 33/216 (15%) 6 y olds; 18/174 (10%) 18 y</td>
<td>&lt;1 y &lt;5 y 6 y 18 y</td>
</tr>
<tr>
<td>Hammitt et al. (2006) [3]</td>
<td>Alaska (USA)</td>
<td>C</td>
<td>377/639 (59%) 5 y olds; 275/2115 (13%) 5–12 y olds; 15/375 (4%) 15 y</td>
<td>&lt;1 y &lt;5 y 5–12 y 15 y</td>
</tr>
<tr>
<td>Henriques Normark et al. (2003) [43]</td>
<td>Sweden</td>
<td>C</td>
<td>246/516 (48%) 1 y olds; 25/375 (6%) 5 y olds; 18/174 (10%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–17 y 18 y</td>
</tr>
<tr>
<td>Hill PC., et al. (2006) [19]</td>
<td>Gambia</td>
<td>C</td>
<td>10/55 (19%) 1 y olds; 51/16 (31%) 5 y olds; 2/38 (5%) 5–15 y olds; 62/123 (51%) 15 y</td>
<td>&lt;1 y &lt;5 y 5–15 y 15 y</td>
</tr>
<tr>
<td>Hussain et al. (2005) [18]</td>
<td>UK</td>
<td>C</td>
<td>87/180 (48%) 1 y olds; 15/71 (21%) 5 y olds; 23/281 (8%) 5–18 y olds; 18/237 (8%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–18 y 18 y</td>
</tr>
<tr>
<td>Inostroza et al. (1998) [44]</td>
<td>Chile</td>
<td>C</td>
<td>340/584 (58%) 1 y olds; 23/109 (21%) 5 y olds; 52/246 (21%) 5–18 y olds; 23/109 (21%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–17 y 18 y</td>
</tr>
<tr>
<td>Kalafot et al. (2008) [45]</td>
<td>Denmark</td>
<td>C</td>
<td>15/59 (25%) 1 y olds; 42/213 (19%) 5 y olds; 47/233 (20%) 5–18 y olds; 47/233 (20%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–18 y 18 y</td>
</tr>
<tr>
<td>Leino et al. (2007) [46]</td>
<td>Finland</td>
<td>C</td>
<td>16/25 (63%) 1 y olds; 107/312 (34%) 5 y olds; 109/553 (20%) 5–10 y olds; 109/553 (20%) 10.5 y</td>
<td>&lt;1 y &lt;5 y 5–10 y 10.5 y</td>
</tr>
<tr>
<td>Leng et al. (2002) [47]</td>
<td>USA</td>
<td>C</td>
<td>32/41 (78%) 1 y olds; 188/342 (55%) 5 y olds; 18/121 (15%) 5–10 y olds; 18/121 (15%) 14 y</td>
<td>&lt;1 y &lt;5 y 5–10 y 14 y</td>
</tr>
<tr>
<td>Lo et al. (2003) [48]</td>
<td>Vietnam</td>
<td>C</td>
<td>16/25 (63%) 1 y olds; 107/312 (34%) 5 y olds; 109/553 (20%) 5–10 y olds; 109/553 (20%) 10.5 y</td>
<td>&lt;1 y &lt;5 y 5–10 y 10.5 y</td>
</tr>
<tr>
<td>Mueller et al. (2012) [49]</td>
<td>Burkina Faso</td>
<td>C</td>
<td>13/25 (52%) 1 y olds; 18/37 (49%) 5 y olds; 28/71 (39%) 5–10 y olds; 28/71 (39%) 14 y</td>
<td>&lt;1 y &lt;5 y 5–10 y 14 y</td>
</tr>
<tr>
<td>Parry et al. (2003) [50]</td>
<td>South Africa</td>
<td>C</td>
<td>212/117 (181%) 1 y olds; 212/117 (181%) 5 y olds; 212/117 (181%) 5–18 y olds; 212/117 (181%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–18 y 18 y</td>
</tr>
<tr>
<td>Regev-Yochay et al. (2004) [51]</td>
<td>Israel</td>
<td>H</td>
<td>38/90 (42%) 1 y olds; 214/379 (56%) 5 y olds; 30/137 (22%) 5–17 y olds; 30/137 (22%) 18 y</td>
<td>&lt;1 y &lt;5 y 5–17 y 18 y</td>
</tr>
<tr>
<td>Regev-Yochay et al. (2012) [52]</td>
<td>Occup. Palest. Terr.</td>
<td>C</td>
<td>188/234 (80%) 1 y olds; 71/248 (29%) 5 y olds; 57/231 (24%) 5–19 y olds; 57/231 (24%) 19 y</td>
<td>&lt;1 y &lt;5 y 5–19 y 19 y</td>
</tr>
<tr>
<td>Sener et al. (1998) [53]</td>
<td>UK</td>
<td>C</td>
<td>49/99 (49%) 1 y olds; 86/172 (50%) 5 y olds; 42/117 (36%) 5–18 y olds; 86/172 (50%) 19 y</td>
<td>&lt;1 y &lt;5 y 5–18 y 19 y</td>
</tr>
<tr>
<td>Turner et al. (2012) [54]</td>
<td>Thailand</td>
<td>C</td>
<td>16/25 (63%) 1 y olds; 107/312 (34%) 5 y olds; 109/553 (20%) 5–10 y olds; 109/553 (20%) 10.5 y</td>
<td>&lt;1 y &lt;5 y 5–10 y 10.5 y</td>
</tr>
<tr>
<td>van Gils E. et al. (2009) [55]</td>
<td>Netherlands</td>
<td>C</td>
<td>214/379 (57%) 1 y olds; 67/379 (22%) 5 y olds; 214/379 (57%) 12 y</td>
<td>&lt;1 y &lt;5 y 12 y</td>
</tr>
</tbody>
</table>

RX: random sample from cross sectional survey, RH: random outpatient sample, SC: school cohort (i.e. all children or staff in a particular school/DCC/class), M: mothers, F: family members, BC: birth cohort. Italicised are the adults samples where adults are parents or staff members looking after the young children included.
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.

We found a positive linear relationship between the proportion of VT serotypes isolated from carriers in children under five and that in 5–17 y 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 5–17 y old and 0.73 (95%CrI 0.06; 1.32) in adult carriers compared to the proportion of VT serotypes in children <5 y. Further details on the model coefficients are provided in Table 3.

We then analysed the relationship between VT and NVT carriage prevalence in <5 y olds and 5–17 y 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 <5 y (Figures 4 and 5). However, given the shift in VT/NVT distribution in older age groups, the prevalence of VT serotypes in 5–17 y olds and adults compared to that in children under five was comparatively lower than that of NVT serotypes (Table 2).

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.

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 <5 y 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 (Th17 cells) [27], and there is evidence that Th17 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 <5 y 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 <5 y 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 5–17 y olds and adults to be 58.1% (95% prediction 38.7–77.4%) and 32.4% (95% prediction 14.8–48.0%) respectively, based on data in children under five. This closely matches the study estimates, with the prevalence in 5–17 y 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, which was seen frequently for the adult age group. However, our subgroup analyses showed no difference in the prediction estimates of the overall carriage prevalence obtained with studies based on random samples of adults compared to studies based on a convenient sample of relatives or carers of the children included.

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 5–17 y 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

| Table 2. Model coefficients for each of the Bayesian linear meta-regression models used. |
|------------------------------------------|------------------------------------------|
| Model                                    | Intercept (I0)                           |
|                                          | median (95% CrI)                         |
|                                          | Slope (I1)                               |
|                                          | median (95% CrI)                         |
| Overall carriage prevalence              |                                         |
| Adults vs. <5 y                          | -0.18 (-0.31; -0.05)                    |
| Adults vs. <1 y                          | -0.36 (-0.49; -0.15)                    |
| 5–17 y vs. <5 y                          | -0.00 (-0.10; 0.10)                     |
| 5–17 y vs. <1 y                          | -0.38 (-0.49; -0.09)                    |
| VT carriage prevalence                   |                                         |
| Adults vs. <5 y                          | -0.06 (-0.17; 0.05)                     |
| 5–17 y vs. <5 y                          | -0.10 (-0.41; 0.20)                     |
| VT proportion among carriers             |                                         |
| Adults vs. <5 y                          | -0.07 (-0.41; 0.33)                     |
| 5–17 y vs. <5 y                          | -0.03 (-0.41; 0.33)                     |

doi:10.1371/journal.pone.0086136.t002
### Table 3. Studies included in the Bayesian meta-regression of VT and NVT distribution, with proportion of VT out of positive samples, by age groups.

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

<sup>1</sup>Maela refugee camp at the Thailand-Myanmar border.  
<sup>2</sup>Gaza strip.  
<sup>*</sup>data marked with (*) are based on approximation. See the Supporting Information section for further details.

doi:10.1371/journal.pone.0086136.t003
many studies the age representation of the 5–17 y 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. There was however good homogeneity in the serotyping methods used in all studies included.

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.

**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.

Supporting Information

Figure S1 Overall carriage prevalence in adults against <5 y olds, by sampling characteristics of the adult population: 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.

References