Determinants of high residual post-PCV13 pneumococcal vaccine type carriage in Blantyre, Malawi: a modelling study.

Additional File 1: methodological details, literature support and complimentary results.

Table of Contents

Ordinary-differential equations (ODE)...........................................................................................................1
Definition of birth rate for a constant population size.........................................................................................2
Expressions for forces of infection.......................................................................................................................2
Expressions for model carriage levels..................................................................................................................3
Markov-chain Monte-Carlo fitting approach........................................................................................................4
Model age-structure at equilibrium.......................................................................................................................4
Literature support for estimated durations of carriage..........................................................................................5
Literature support for estimated vaccine efficacy against carriage....................................................................5
Literature support and model sensitivity to different transmission matrices.......................................................6
Literature review of carriage reduction in time.......................................................................................................9
Literature review of pre-vaccination carriage levels.............................................................................................10
Extra results for sensitivity of vaccine impact projections..................................................................................11
Observational study data.......................................................................................................................................12
References..........................................................................................................................................................13

Ordinary-differential equations (ODE)

Equations for the vaccinated (equations 1-14) and unvaccinated (equations 15-28) age-groups, with
model diagram in Figure 1 of main text:

\[
\begin{align*}
\frac{dS_0}{dt} &= b(1 - \rho) - \lambda_0S_0 + \gamma_0c_0 - a_0S_0 - \mu_0S_0 \\
\frac{dC_0}{dt} &= \lambda_0S_0 - \gamma_0c_0 - a_0c_0 - \mu_0c_0 \\
\frac{dS_1}{dt} &= a_0S_0 - \lambda_1S_1 + \gamma_1c_1 - a_1S_1 - \mu_1S_1 \\
\frac{dC_1}{dt} &= a_0C_0 + \lambda_1S_1 - \gamma_1c_1 - a_1C_1 - \mu_1C_1 \\
\frac{dS_2}{dt} &= a_1S_1 - \lambda_2S_2 + \gamma_2c_2 - a_2S_2 - \mu_2S_2 \\
\frac{dC_2}{dt} &= a_1C_1 + \lambda_2S_2 - \gamma_2c_2 - a_2C_2 - \mu_2C_2 \\
\frac{dS_{3-5}}{dt} &= a_2S_2 - \lambda_{3-5}S_{3-5} + \gamma_{3-5}C_{3-5} - a_{3-5}S_{3-5} - \mu_{3-5}S_{3-5} \\
\frac{dC_{3-5}}{dt} &= a_2C_2 + \lambda_{3-5}S_{3-5} - \gamma_{3-5}C_{3-5} - a_{3-5}C_{3-5} - \mu_{3-5}C_{3-5} \\
\frac{dS_{6-7}}{dt} &= a_{3-5}S_{3-5} - \lambda_{6-7}S_{6-7} + \gamma_{6-7}C_{6-7} - a_{6-7}S_{6-7} - \mu_{6-7}S_{6-7} \\
\frac{dC_{6-7}}{dt} &= a_{3-5}C_{3-5} + \lambda_{6-7}S_{6-7} - \gamma_{6-7}C_{6-7} - a_{6-7}C_{6-7} - \mu_{6-7}C_{6-7} \\
\frac{dS_{8-9}}{dt} &= a_{6-7}S_{6-7} - \lambda_{8-9}S_{8-9} + \gamma_{8-9}C_{8-9} - a_{8-9}S_{8-9} - \mu_{8-9}S_{8-9} \\
\frac{dC_{8-9}}{dt} &= a_{6-7}C_{6-7} + \lambda_{8-9}S_{8-9} - \gamma_{8-9}C_{8-9} - a_{8-9}C_{8-9} - \mu_{8-9}C_{8-9} \\
\frac{dS_{10+}}{dt} &= a_{8-9}S_{8-9} - \lambda_{10+}S_{10+} + \gamma_{8-9}C_{10+} - a_{10+}S_{10+} - \mu_{10+}S_{10+} \\
\frac{dC_{10+}}{dt} &= a_{8-9}C_{8-9} + \lambda_{10+}S_{10+} - \gamma_{8-9}C_{10+} - a_{10+}C_{10+} - \mu_{10+}C_{10+}
\end{align*}
\]
Definition of birth rate for a constant population size

Given that the total population size is kept constant, the birth rate is a composite expression of deaths across ages:

\[
b = \mu_0(S_0 + C_0 + S_0^0 + C_0^0) + \\
\mu_1(S_1 + C_1 + S_1^0 + C_1^0) + \\
\mu_2(S_2 + C_2 + S_2^0 + C_2^0) + \\
\mu_3(S_3 + C_3 + S_3^0 + C_3^0) + \\
\mu_4(S_4 + C_4 + S_4^0 + C_4^0) + \\
\mu_5(S_5 + C_5 + S_5^0 + C_5^0) + \\
\mu_6(S_6 + C_6 + S_6^0 + C_6^0) + \\
\mu_7(S_7 + C_7 + S_7^0 + C_7^0) + \\
\mu_8(S_8 + C_8 + S_8^0 + C_8^0) + \\
\mu_9(S_9 + C_9 + S_9^0 + C_9^0) + \\
\mu_{10}(S_{10} + C_{10} + S_{10}^0 + C_{10}^0)
\] (29)

Expressions for forces of infection

The forces of infection of the different model variants (see Figure S1) are defined as follows.

model 0-5 years + Assortative

\[
\lambda_0 = \theta C_a_0 + \theta C_a_1 + \theta C_a_2 + \theta C_a_{3-5} + \beta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (30)

\[
\lambda_1 = \theta C_a_0 + \theta C_a_1 + \theta C_a_2 + \theta C_a_{3-5} + \beta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (31)

\[
\lambda_2 = \theta C_a_0 + \theta C_a_1 + \theta C_a_2 + \theta C_a_{3-5} + \beta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (32)

\[
\lambda_{3-5} = \theta C_a_0 + \theta C_a_1 + \theta C_a_2 + \theta C_a_{3-5} + \beta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (33)

\[
\lambda_{6-7} = \beta C_a_0 + \beta C_a_1 + \beta C_a_2 + \beta C_a_{3-5} + \theta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (34)

\[
\lambda_{8-9} = \beta C_a_0 + \beta C_a_1 + \beta C_a_2 + \beta C_a_{3-5} + \beta C_a_{0-7} + \theta C_a_{8-9} + \beta C_a_{10+}
\] (35)

\[
\lambda_{10+} = \beta C_a_0 + \beta C_a_1 + \beta C_a_2 + \beta C_a_{3-5} + \beta C_a_{0-7} + \beta C_a_{8-9} + \beta C_a_{10+}
\] (36)

model 1-5 years + Assortative
\begin{equation}
\lambda_0 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{37}
\end{equation}
\begin{equation}
\lambda_1 = \beta C a_0 + \theta C a_1 + \theta C a_2 + \theta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{38}
\end{equation}
\begin{equation}
\lambda_2 = \beta C a_0 + \theta C a_1 + \theta C a_2 + \theta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{39}
\end{equation}
\begin{equation}
\lambda_{3-5} = \beta C a_0 + \theta C a_1 + \theta C a_2 + \theta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{40}
\end{equation}
\begin{equation}
\lambda_{6-7} = \beta C a_0 + \theta C a_1 + \theta C a_2 + \theta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{41}
\end{equation}
\begin{equation}
\lambda_{8-9} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \theta C a_{8-9} + \beta C a_{10+} \tag{42}
\end{equation}
\begin{equation}
\lambda_{10+} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{43}
\end{equation}

model 2-5 years + Assortative

\begin{equation}
\lambda_0 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{44}
\end{equation}
\begin{equation}
\lambda_1 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{45}
\end{equation}
\begin{equation}
\lambda_2 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{46}
\end{equation}
\begin{equation}
\lambda_{3-5} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{47}
\end{equation}
\begin{equation}
\lambda_{6-7} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{48}
\end{equation}
\begin{equation}
\lambda_{8-9} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{49}
\end{equation}
\begin{equation}
\lambda_{10+} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{50}
\end{equation}

model Assortative

\begin{equation}
\lambda_0 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{51}
\end{equation}
\begin{equation}
\lambda_1 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{52}
\end{equation}
\begin{equation}
\lambda_2 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{53}
\end{equation}
\begin{equation}
\lambda_{3-5} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{54}
\end{equation}
\begin{equation}
\lambda_{6-7} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{55}
\end{equation}
\begin{equation}
\lambda_{8-9} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \theta C a_{8-9} + \beta C a_{10+} \tag{56}
\end{equation}
\begin{equation}
\lambda_{10+} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{57}
\end{equation}

model Homogeneous

\begin{equation}
\lambda_0 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{58}
\end{equation}
\begin{equation}
\lambda_1 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{59}
\end{equation}
\begin{equation}
\lambda_2 = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{60}
\end{equation}
\begin{equation}
\lambda_{3-5} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{61}
\end{equation}
\begin{equation}
\lambda_{6-7} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{62}
\end{equation}
\begin{equation}
\lambda_{8-9} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{63}
\end{equation}
\begin{equation}
\lambda_{10+} = \beta C a_0 + \beta C a_1 + \beta C a_2 + \beta C a_{3-5} + \beta C a_{6-7} + \beta C a_{8-9} + \beta C a_{10+} \tag{64}
\end{equation}

Expressions for model carriage levels

Levels of carriage in the ODE model are calculated as the proportion of individuals within an age-group, with a particular vaccine status, that are carriers at the time of the surveys. For example:

\begin{align*}
C_{a_0} &= C_0 + C_{a_0}^\nu \tag{65} \\
C_{a_1} &= C_1 + C_{a_1}^\nu \tag{66} \\
C_{a_2} &= C_2 + C_{a_2}^\nu \tag{67} \\
C_{a_{3-5}} &= C_{3-5} + C_{a_{3-5}}^\nu \tag{68} \\
C_{a_{6-7}} &= C_{6-7} + C_{a_{6-7}}^\nu \tag{69} \\
C_{a_{8-9}} &= C_{8-9} + C_{a_{8-9}}^\nu \tag{70} \\
C_{a_{10+}} &= C_{10+} + C_{a_{10+}}^\nu \tag{71}
\end{align*}
Markov-chain Monte-Carlo fitting approach

We used a Bayesian Markov chain Monte Carlo (bMCMC) approach, developed and used by us in other modelling studies [1]–[3]. The proposal distributions \((q)\) of each parameter are defined as Gaussian (symmetric), effectively implementing a random walk Metropolis kernel. We define our acceptance probability \(\alpha\) of a parameter set \(\Theta^*\) as:

\[
\alpha = \min\{1, \frac{\pi(y | \Theta^*) p(\Theta^*) q(\Theta^* | \Theta^o)}{\pi(y | \Theta^o) p(\Theta^o) q(\Theta^o | \Theta^*)}\}
\]

where \(\Theta^*\) and \(\Theta^o\) are the proposed and current (accepted) parameter sets (respectively); \(\Pi(y \mid \Theta^*)\) and \(\Pi(y \mid \Theta^o)\) are the likelihoods of the observed data \((y, \text{Table S7})\) under the ODE output given the parameter sets; \(p(\Theta^o)\) and \(p(\Theta^*)\) are the prior-related probabilities given each parameter set.

For simplicity and because all fitted variables are proportions, the likelihoods \(\pi\) were calculated as the product of conditional Gaussian probabilities \((f_{\Theta}(y_i))\). The likelihood is the product the conditional probabilities of all variables, and can be formally expressed as:

\[
\pi(y \mid \Theta) = \prod_{i=1}^{N} [f_{\Theta}(y_i)]
\]

We used Gaussian probabilities instead of Binomial for mathematical and computational convenience. It should be noted, however, that the Binomial distribution is well approximated by the Gaussian distribution under the common rule of thumb of \(n^* \min(p, 1-p) > 5\) \((n=\text{sample size, } p=\text{carriage level in an age-group})\), which is not violated in our observational dataset [4].

Model age-structure at equilibrium

The age-groups used in the modelling framework were based on the groups for which carriage data was collected in the rolling prospective nasopharyngeal carriage surveys [4]. To confirm that ageing was robust in the ODE model, we compared the model’s proportions of the population in each age-group (at equilibrium, before vaccination) to the known proportions for Malawi (country’s age pyramid for the year of 2019 available at https://www.populationpyramid.net/malawi/2019/).

Because the age pyramid groups under 10 years of age were different from the model’s, we first compared proportions below and above 10 years of age: we obtained mean model proportions of 0.150 and 0.849 versus age pyramid proportions of 0.158 and 0.842 (for <10 and >=10 years, respectively). We then attempted to approximate the age-grouping under 10 years of age; using the pyramid’s age-groups 0-4 and 5-9 years. For this, we aggregated the model’s proportions with all 0, 1, 2, and two-thirds of 3-5 years old, as well as aggregated the proportions with one-third of 3-5, and all 6-7 and 8-9 years old. The model’s proportions were 0.078 and 0.072, compared to the pyramid’s 0.084 and 0.074 (for 0-4 and 5-9 years, respectively). The proportions between the model and the age pyramid were very similar, thus validating the ageing implemented in the ODE model.
Literature support for estimated durations of carriage

While we base our priors on the study by Hogberg and colleagues [5], there is vast support in the literature for both the use of similar values and the key assumption of a decrease in duration of carriage with age. We summarize this support in Table S1 below.

<table>
<thead>
<tr>
<th>Age-group</th>
<th>Duration</th>
<th>Country</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1 years</td>
<td>71</td>
<td>UK</td>
<td>[11]–[13]</td>
</tr>
<tr>
<td>2-4 years</td>
<td>28</td>
<td>UK</td>
<td>[11]–[13]</td>
</tr>
<tr>
<td>5-17 years</td>
<td>18</td>
<td>UK</td>
<td>[11]–[13]</td>
</tr>
<tr>
<td>18+ years</td>
<td>17</td>
<td>UK</td>
<td>[11]–[13]</td>
</tr>
<tr>
<td>0-1 years</td>
<td>30</td>
<td>Sweden</td>
<td>[14]</td>
</tr>
<tr>
<td>1-4 years</td>
<td>21</td>
<td>Sweden</td>
<td>[14]</td>
</tr>
<tr>
<td>5-6 years</td>
<td>13</td>
<td>Sweden</td>
<td>[14]</td>
</tr>
<tr>
<td>7-18 years</td>
<td>15</td>
<td>Sweden</td>
<td>[6]</td>
</tr>
<tr>
<td>18+ years</td>
<td>14</td>
<td>Sweden</td>
<td>[6]</td>
</tr>
<tr>
<td>0-2 years</td>
<td>60*</td>
<td>Thailand &amp; Myanmar</td>
<td>[7]</td>
</tr>
<tr>
<td>Mothers</td>
<td>31*</td>
<td>Thailand &amp; Myanmar</td>
<td>[7]</td>
</tr>
<tr>
<td>6-11 months</td>
<td>48.5*</td>
<td>Kenya</td>
<td>[8]</td>
</tr>
<tr>
<td>3-59 months</td>
<td>31.3*</td>
<td>Kenya</td>
<td>[8]</td>
</tr>
<tr>
<td>0-7 years</td>
<td>44*</td>
<td>Finland</td>
<td>[9]</td>
</tr>
<tr>
<td>7+ years</td>
<td>36</td>
<td>Finland</td>
<td>[9]</td>
</tr>
<tr>
<td>0-5 years</td>
<td>42.6*</td>
<td>Bangladesh</td>
<td>[10]</td>
</tr>
<tr>
<td>5+ years</td>
<td>38</td>
<td>Bangladesh</td>
<td>[10]</td>
</tr>
</tbody>
</table>

Table S1 – Support for decrease of duration of carriage with age. Mean values are presented and * marks values for which 95% CI are available in the original study. Grey and white backgrounds are a visual cue for different countries.

Literature support for estimated vaccine efficacy against carriage

The posterior for vaccine efficacy against carriage (ζ) was shown in the main text to be in the range of estimates obtained from other studies. Table S2 presents the values reported in such studies.

<table>
<thead>
<tr>
<th>VE</th>
<th>Vaccine</th>
<th>Ref.</th>
<th>Country / Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>56% (95% CI 41 - 72)</td>
<td>PCV10</td>
<td>[15]</td>
<td>Kenya</td>
</tr>
<tr>
<td>66% (95% CI 38 - 82)</td>
<td>PCV10</td>
<td>[16]</td>
<td>Kenya</td>
</tr>
<tr>
<td>63% (95% CI 53 - 73)</td>
<td>Meta</td>
<td>[17]</td>
<td>Meta-analysis</td>
</tr>
<tr>
<td>78% (95% CI 64 - 92)</td>
<td>PCV13</td>
<td>[18]</td>
<td>Vietnam</td>
</tr>
<tr>
<td>59% (95% CI 49 - 68)</td>
<td>PCV10</td>
<td>[19]</td>
<td>Kenya</td>
</tr>
</tbody>
</table>

Table S2 – Support for estimated vaccine efficacy. Mean values, 95% CI, vaccine, reference and country are presented for each of the studies. These quantifications were reported in the original articles (were not estimated by us on reported data).
Literature support and model sensitivity to different transmission matrices

We formulated five different transmission matrices (Figure S1). One matrix assumed that transmission potential between age-groups is homogeneous (Figure S1e), while the other four assumed variations of inhomogeneous transmission potential between the groups (Figure S1a-d).

Figure S1: Transmission matrices used for sensitivity of the model. Subplots (a) to (e) present the five transmission matrices used in five independent fitting exercises of the model to the survey data described in the main text. The names of each matrix (and model) is above each subplot. The \( \beta \) and \( \theta \) letters refer to the transmission coefficients described in the main text. The coefficient \( \theta \) is in principle larger than \( \beta \), but this relationship is not explicitly forced in the fitting approach.

The assumption, observation or estimation that pneumococcal transmission is inhomogeneous across ages is common in the literature. For instance, Ojal et al. have estimated that the probabilities of acquisition (colonization) across ages in Kilifi, Kenya after the introduction of PCV10, concluding that the probability of the age-group 1-5 years old is about 1.25 and 5.7 times higher than in the age-groups 6-14 and 15-20 years old [20]. For Finland [21], it has been estimated that around 44% of colonization events have a child <7 years of age as their source and that 18% of transmission events occurred within this age group; and in Italy, the odds ratio for colonization in children aged <2 years after PCV13 introduction was still high at 3.75 [22]. In a different study by Adetifa et al., carriage was 1.4 times higher in children <4 years of age compared to the next age-group 5-14 years of age [23]. Still in Africa, Uganda, Polain de Waroux et al. found that the age-group 5-14 years had the highest frequency of contacts, and a strong age-assortative pattern was seen across age-groups [24]. The latter effect has been reported in several countries including Uganda, Kenya, Finland, USA, and Great Britain [21], [24]–[27]. The epidemiological importance of considering higher efficiency and/or contact numbers between and within younger age-groups is not restricted to the pneumococcus, and is often reported for other pathogens as well [28].

After fitting the model to the survey data in five independent exercises (each with one of the matrices in Figure S1), we compared the frameworks using leave-one-out cross-validation (LOO) and widely applicable information criterion (WAIC) measures. Comparisons were done based on (1) the estimated posteriors versus priors, and (2) posterior model carriage levels versus observed surveys 1 – 7. The priors included were the carriage duration by Hogberg and colleagues [5] (Table S1, Figure 1b) and the vaccine efficacy against carriage (Table S2 and Figure 2c). LOO and WAIC are methods that allow estimating pointwise out-of-sample prediction accuracy from the five models using the log-likelihood evaluated at simulations from the parameter values of the estimated
posteriors [29]. We used the loo R-package [30] to calculate model weights: 1) WAIC weights, 2) Pseudo-BMA weights without Bayesian bootstrap, 3) Pseudo-BMA + weights with Bayesian bootstrap, and 4) Bayesian stacking weights [31]. All of these measures vary between 0 and 1, with the sum between models adding up to 1; the highest the weight, the more favoured a model is compared to the rest. That is, a higher weight points to a more balanced response from the model, both in terms of reproducing the survey data but also in respecting the priors for duration of carriage and the literature knowledge on vaccine efficacy against carriage. The results of model comparison are in Table S3.

<table>
<thead>
<tr>
<th>model / matrix</th>
<th>1) WAIC</th>
<th>2) Pseudo-BMA</th>
<th>3) Pseudo-BMA with BS</th>
<th>4) Bayesian stacking</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5 years + Assortative</td>
<td>0.69</td>
<td>0.99</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>1-5 years + Assortative</td>
<td>0.31</td>
<td>0.01</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>2-5 years + Assortative</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Assortative</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Homogeneous</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table S3 – LOO and WAIC weights between model frameworks.

The best score was for the model with the 0-5 years + Assortative transmission matrix (Figure 1a). This is therefore the model presented in the main text. In this model, the transmission potential (θ) among individuals 0-5 years old, and among 7-6 and 8-9 years old is different (larger, Figure 2d) than among other age-groups (β). The second best scored model (1-5 years + Assortative, Figure 1b) was the one using the most similar matrix to 0-5 years + Assortative, but which critically did not consider individuals of age 0 years old (i.e. <1 years). For completion, we present the general output of the 1-5 years + Assortative model in Figure S2.

<table>
<thead>
<tr>
<th>age-group</th>
<th>Carriage reduction after 1 year</th>
<th>age-group</th>
<th>Carriage reduction after 10 years</th>
<th>age-group</th>
<th>Posterior adjustment to priors of carriage duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>48.2 (38.7-57.1)</td>
<td>0</td>
<td>79.8 (67.9-87.7)</td>
<td>0-2</td>
<td>-0.68</td>
</tr>
<tr>
<td>1</td>
<td>26.87 (23.1-29.7)</td>
<td>1</td>
<td>75.1 (65.2-81.7)</td>
<td>3-5</td>
<td>+0.58</td>
</tr>
<tr>
<td>2</td>
<td>14.9 (13.4-15.9)</td>
<td>2</td>
<td>74.3 (64.4-80.7)</td>
<td>6-7</td>
<td>+0.53</td>
</tr>
<tr>
<td>3-5</td>
<td>7.3 (6.6-7.8)</td>
<td>3-5</td>
<td>74.6 (64.9-80.8)</td>
<td>8+</td>
<td>-1.71</td>
</tr>
<tr>
<td>0-5</td>
<td>16.5 (14.9-17.1)</td>
<td>0-9</td>
<td>75.1 (64.5-81.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-9</td>
<td>6.1 (5.6-6.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table S4 – Posteriors for the model 1-5 years + Assortative. Ranges are the 95% CI.

The 1-5 years + Assortative model presents differences in output and capacity to fit the data. For instance, large differences in the posterior for vaccine efficacy against carriage and the estimation of pre-vaccination VT carriage among 0-5 years old can be visually identified comparing Figure 2c,
Figure 5a and Figure S2. For completion, we include here the 1-5 years + Assortative model full range of results (Table S4) presented in the main text for the 0-5 years + Assortative.

Although we defined the model transmission matrices similarly to matrices reported in previous studies, it is likely that further important heterogeneities exist between age-groups. To assess this, however, would require the addition of several parameters to the system, which in turn would increase the complexity and uncertainty of the fitting exercise. It is also the case that without further empirical support, it is difficult to plan and develop this new parametrization, as the required details on which age-group have specific lower / higher transmission coefficients in the particular context of the Blantyre cohort is largely unknown.

The independent contribution of either mixing (e.g. contact frequency) or risk (e.g. efficiency of contact) to the force of infection of each age-group can not be disentangled in our modelling framework. Nonetheless, we note that both higher mixing and risk of infection are supported in our study. For the first, as described above, we resort to literature reports on particular mixing patterns between age groups. For the second, we note that a higher proportion of children 3-6 years of age lived in houses with some lower quality infrastructure and had greater reliance on shared communal water sources in the cohort of our observational study [4].
Figure S2: Model fit and estimated posteriors when using a framework that uses the 1-5 years + Assortative transmission matrix. (a) Model fit to carriage data from the observational study for age-groups: vaccinated 2 years old (red), vaccinated 3-5 years old (purple), unvaccinated 6-7 years old (green) and unvaccinated 8-9 years old (orange). The survey data is represented with means as full circles, the model output with means as full squares; the whiskers are the 95% CI. (b) Priors (lines) and estimated posterior distributions (shaded areas) for duration of carriage per age-group. (c) Estimated posterior for vaccine efficacy (and 95% CI) against vaccine-type carriage (red) in the context of estimations from other studies (in legend, Table S2). (d) Estimated posterior for the transmission coefficient \( \beta \). (e) Projected reduction in carriage relative to the pre-vaccination era for age-groups 0 years (magenta), 1 year (blue), 2 years (yellow) and 3-5 years (purple) old. (f) Projected reduction in carriage relative to the pre-vaccination era for aggregated age-groups 0-5 years (green) and 6-9 years (red) old (with corresponding 95% CIs). The shaded areas are: yellow for the post-vaccination period with no carriage data, white for the post-vaccination period with survey carriage data, and grey for the post-vaccination projected period up to 10 years. (g) Estimated pre-vaccination vaccine-type carriage (and 95% CI) for the age-group 0-5 years of age (red) in the context of carriage levels reported in other studies (in legend, Table S6). (h) The baseline transmission coefficient (\( \beta \)) is varied by considering the 70%, 60%, 50%, 40%, 30%, 20%, and 10% lower, and 10%, 20% higher transmission than the estimated for Blantyre (Malawi, \( \beta_{\text{Malawi}} \)) when fitting the observational study (e.g. 10% lower is 0.9*\( \beta_{\text{Malawi}} \)). The impact projections for the age-group 0-5 years old using the \( \beta \) estimated for Blantyre (Malawi) are presented by the dashed line (as in Figure 3b). The symbols and whiskers are measures of reported impact (carriage reduction) and 95% CIs for several published studies (in legend, Table S5). (a,b,c,d,e,f,g,h) Solutions presented are obtained from sampling 100,000 parameter values from posteriors and simulating the dynamic model.

Literature review of carriage reduction in time

The pre- and post-vaccination proportions of vaccine-type carriage were used to calculate a Relative Risk (RR) and its corresponding 95% CI (under asymptotic normality). Using the expression \( \frac{\text{carriage(post)} - \text{carriage(pre)}}{\text{carriage(post)}} = 1 - \text{carriage(pre)} / \text{carriage(post)} = 1 - \text{RR} \) yielded a corresponding 95% CI for the relative change in carriage (Table S5) for particular time points after vaccination, age-groups, vaccines and countries. These values are used in figures of the main text and other supplementary figures comparing results from Blantyre (Malawi) with other countries.

<table>
<thead>
<tr>
<th>Percent Reduction</th>
<th>Time</th>
<th>Age</th>
<th>Vaccine</th>
<th>Ref.</th>
<th>Country / Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.13% (95% CI 70.31 - 86.71)</td>
<td>3</td>
<td>11 m</td>
<td>PCV7</td>
<td>[36]</td>
<td>Netherlands</td>
</tr>
<tr>
<td>69.33% (95% CI 46.65 - 82.37)</td>
<td>4</td>
<td>6–23 m</td>
<td>PCV7</td>
<td>[43]</td>
<td>USA</td>
</tr>
<tr>
<td>34.84% (95% CI 24.52 - 43.74)</td>
<td>2</td>
<td>6–24 m</td>
<td>PCV7</td>
<td>[50]</td>
<td>France</td>
</tr>
<tr>
<td>52.62% (95% CI 43.6 - 60.2)</td>
<td>3</td>
<td>6–24 m</td>
<td>PCV7</td>
<td>[50]</td>
<td>France</td>
</tr>
<tr>
<td>60.88% (95% CI 53.27 - 67.26)</td>
<td>4</td>
<td>6–24 m</td>
<td>PCV7</td>
<td>[50]</td>
<td>France</td>
</tr>
<tr>
<td>49.63% (95% CI 33.99 - 61.57)</td>
<td>1</td>
<td>6–24 m</td>
<td>PCV13</td>
<td>[51]</td>
<td>France</td>
</tr>
<tr>
<td>95.31% (95% CI 93.44 - 96.65)</td>
<td>5</td>
<td>&lt;2 y</td>
<td>PCV13</td>
<td>[52]</td>
<td>China</td>
</tr>
<tr>
<td>88.05% (95% CI 79.64 - 92.99)</td>
<td>3</td>
<td>2 y</td>
<td>PCV7</td>
<td>[36]</td>
<td>Netherlands</td>
</tr>
<tr>
<td>77.17% (95% CI 53.76 - 88.73)</td>
<td>1</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[32]</td>
<td>USA</td>
</tr>
<tr>
<td>78.91% (95% CI 73.16 - 83.42)</td>
<td>5</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[33]</td>
<td>Portugal</td>
</tr>
<tr>
<td>61.44% (95% CI 46.63 - 72.14)</td>
<td>1</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[34]</td>
<td>USA</td>
</tr>
<tr>
<td>80.32% (95% CI 69.28 - 87.39)</td>
<td>2</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[34]</td>
<td>USA</td>
</tr>
<tr>
<td>91.41% (95% CI 83.65 - 95.49)</td>
<td>3</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[34]</td>
<td>USA</td>
</tr>
<tr>
<td>32.24% (95% CI 25.93 - 38.02)</td>
<td>1</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[35]</td>
<td>Canada</td>
</tr>
</tbody>
</table>
Table S5 – Percent reduction in carriage for several studies. Mean, 95% CI, age-group, time since vaccination, vaccine, reference and country are presented for each study. See text for mean and CI calculation.

**Literature review of pre-vaccination carriage levels**

Table S6 presents a literature review of pre-vaccination carriage levels for different groups of (vaccine) serotypes in different epidemiological contexts. These values are used in figures of the main text comparing results from Blantyre (Malawi) with other countries.

<table>
<thead>
<tr>
<th>Carriage in %</th>
<th>Age</th>
<th>Vaccine types</th>
<th>Ref.</th>
<th>Country / Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>53.80% (95% CI 46.02 - 61.44)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[39]</td>
<td>USA</td>
</tr>
<tr>
<td>27% (95% CI 23 - 32)</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[18]</td>
<td>Vietnam</td>
</tr>
<tr>
<td>67.87% (95% CI 63.44 - 72.07)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[35]</td>
<td>Canada</td>
</tr>
<tr>
<td>24.7% (95% CI 24.7 - 24.7)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[53]</td>
<td>Gambia</td>
</tr>
<tr>
<td>53.13% (95% CI 49.40 - 56.84)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[33]</td>
<td>Portugal</td>
</tr>
<tr>
<td>55.43% (95% CI 50.26 - 60.52)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[34]</td>
<td>USA</td>
</tr>
<tr>
<td>19.04% (95% CI 10.24 - 30.90)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[55]</td>
<td>UK</td>
</tr>
<tr>
<td>29.03% (95% CI 21.23 - 37.86)</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[37]</td>
<td>USA</td>
</tr>
<tr>
<td>15.9% (95% CI 13.06 - 18.98)</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[38]</td>
<td>Italy</td>
</tr>
<tr>
<td>46.8% (95% CI 46.8 - 46.8)</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[53]</td>
<td>Gambia</td>
</tr>
<tr>
<td>46.4% (95% CI 41.26 - 51.59)</td>
<td>&lt;5 y</td>
<td>PCV7</td>
<td>[23]</td>
<td>Nigeria</td>
</tr>
<tr>
<td>70.4% (95% CI 65.49 - 74.97)</td>
<td>&lt;5 y</td>
<td>PCV13</td>
<td>[23]</td>
<td>Nigeria</td>
</tr>
</tbody>
</table>
Table S6 – Pre-vaccination carriage levels for several studies. Mean, 95% CI, age-group, vaccine types (serotypes of particular PCV vaccine), reference and country are presented for each study.

Extra results for sensitivity of vaccine impact projections

In Figures S3 and S4 we present extra results on the sensitivity of vaccine impact projections relative to baseline transmission, as an extension to main Figure 5 (and thus using the transmission matrix of Figure S1a. Figure S3 is based on all age-groups under the age of 10 years old as projected by the model (for comparison, the age-groups of main Figure 5 are presented again). Figure S4 presents two age classes for which we found empirical reports on carriage decrease post-vaccination, as an extension to main Figure 5.

Figure S3: Extra results for sensitivity of impact projections to baseline transmission (model age-groups under the age of 10 years).

Figure 5 of the main text presents results for age-group 0-5 years of age. This figure presents results for all age-groups. The baseline transmission is varied by considering the 70%, 60%, 50%, 40%, 30%, 20%, 10% lower, and 10%, 20%, 30% higher relative change to the original values estimated for \( \beta \) and \( \theta \) when fitting the observational study’s data (e.g. 40% lower means 0.6*\( \beta \) and 0.6*\( \theta \)). The impact projections using \( \beta \) and \( \theta \) of Blantyre (Malawi) are presented with the shaded line (same as in main text). For visual purposes, only the means are shown, obtained from simulations sampling 100,000 parameter values from posteriors.
Figure S4: Extra results for sensitivity of impact projections to baseline transmission (two model age-groups with empirical observations). Figure 5 of the main text presents results for age-group 0-5 years of age including reported data. This figure presents results for two other age-groups for which reports were available. The baseline transmission is varied by considering the 70%, 60%, 50%, 40%, 30%, 20%, 10% lower, and 10%, 20%, 30% higher relative change to the original values estimated for β and θ when fitting the observational study’s data (e.g. 40% lower means 0.6*β and 0.6*θ). The impact projections using β and θ of Blantyre (Malawi) are presented with the shaded line (same as in main text). For visual purposes, only the means are shown, obtained from simulations sampling 100,000 parameter values from posteriors.

Observational study data

Table S7 keeps the carriage levels observed in the observational surveys and used to fit the model.

<table>
<thead>
<tr>
<th>survey \ age-group</th>
<th>Vacc. 2 years</th>
<th>Vacc. 3-5 years</th>
<th>Unvacc. 6-7 years</th>
<th>Unvacc. 8-9 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey 1</td>
<td>No data</td>
<td>0.199 (0.0236)</td>
<td>0.283 (0.0619)</td>
<td>0.208 (0.0414)</td>
</tr>
<tr>
<td>Survey 2</td>
<td>No data</td>
<td>0.205 (0.0232)</td>
<td>0.237 (0.0488)</td>
<td>0.145 (0.0424)</td>
</tr>
<tr>
<td>Survey 3</td>
<td>No data</td>
<td>0.208 (0.0214)</td>
<td>0.212 (0.0411)</td>
<td>0.19 (0.0428)</td>
</tr>
<tr>
<td>Survey 4</td>
<td>0.218 (0.0371)</td>
<td>0.175 (0.0196)</td>
<td>0.176 (0.0413)</td>
<td>0.123 (0.0366)</td>
</tr>
<tr>
<td>Survey 5</td>
<td>0.184 (0.0363)</td>
<td>0.194 (0.0205)</td>
<td>0.125 (0.0585)</td>
<td>0.0851 (0.0407)</td>
</tr>
<tr>
<td>Survey 6</td>
<td>0.191 (0.0337)</td>
<td>0.15 (0.0183)</td>
<td>0.132 (0.0465)</td>
<td>0.112 (0.0353)</td>
</tr>
<tr>
<td>Survey 7</td>
<td>0.151 (0.0319)</td>
<td>0.167 (0.0207)</td>
<td>0.2 (0.179)</td>
<td>0.162 (0.0341)</td>
</tr>
</tbody>
</table>

Table S7 – Carriage data from the observational study (mean and standard error)
<table>
<thead>
<tr>
<th>References</th>
</tr>
</thead>
</table>


