Supplementary Material

S1 Supplementary data

S1.1 Inter-individual contact data

Data on age-specific contact rates were available from a study of self-reported contacts in France [1], comparable to the POLYMOD study in other European countries [2]. The corresponding contact matrix, denoted by $c = (c_{ij})$, is plotted in Fig. S1 and was used in all model simulations. In a sensitivity analysis, we also tested the POLYMOD contact matrix from Great Britain (see Fig. S7).

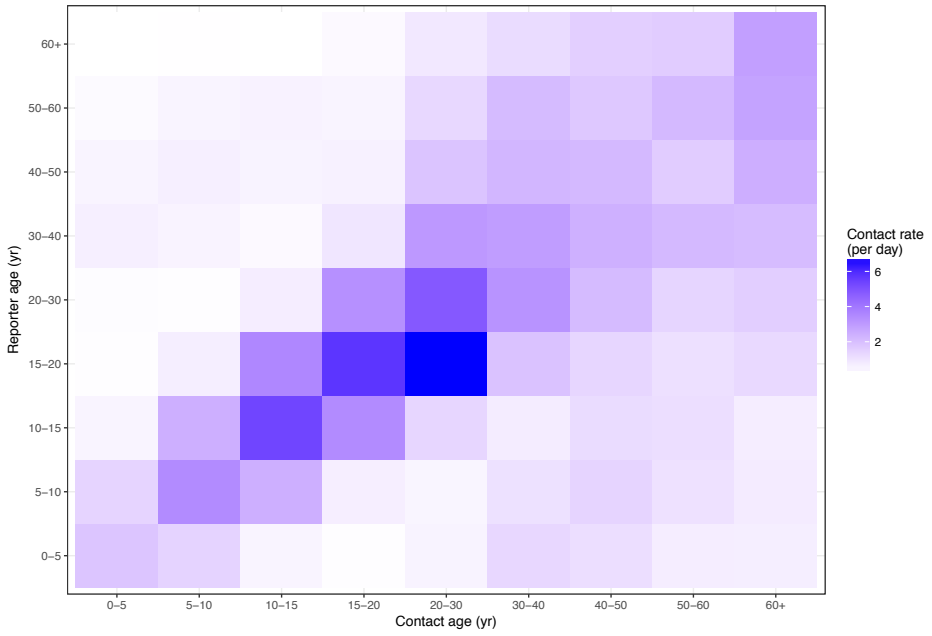

Figure S1: **Matrix of age-specific contact rates in France**. Data are from Ref. [1].

S1.2 Demographic data

The annual number of births and the annual age-specific estimates of population sizes were available from the French National Institute of Statistics and Economic Studies[1]. These data were interpolated using cubic smoothing splines (with 10 degrees of freedom) to calculate the weekly birth rate, $B(t)$, and the population size in every age group i, $N_i(t)$. These smoothed quantities were then used to estimate age-specific migration rates $\mu_i(t)$, according to the equations:

[1] https://www.insee.fr/fr/statistiques, downloaded on 23 December 2016.

$$\begin{aligned}\frac{\mathrm{d}N_1}{\mathrm{d}t} &= B(t) - [\delta_1 + \mu_1(t)]N_i(t) \\ \frac{\mathrm{d}N_{i\geq 2}}{\mathrm{d}t} &= \delta_{i-1}N_{i-1}(t) - [\delta_i + \mu_i(t)]N_i(t)\end{aligned}$$

where δ_i is the aging rate (per wk) in age group i. The migrations rates were incorporated as covariates into all the models, so that the simulated population sizes approximately matched the observed population sizes.

S2 Supplementary methods

S2.1 Principal component analysis of meteorological data

Daily meteorological records from nine weather stations in France were provided by Météo-France, the French national meteorological service. The weather stations were located near the most populated cities across France: Paris (Ile-de-France region), Rennes and Nantes (Northwest region), Lille and Strasbourg (Northeast region), Lyon and Marseille (Southeast region), and Toulouse and Bordeaux (Southwest region). As explained in the main text, four meteorological variables were considered and country-level weekly time series were constructed by averaging each variable temporally and spatially. The resulting time series are displayed in the top four panels of Fig. S3.

Because these variables were markedly correlated, we performed a principal component analysis (PCA) to summarize them [3]. As shown in Fig. S2, the first two components, denoted by $M_1(t)$ and $M_2(t)$, captured about 76% and 21% of the variation in the meteorological data. Inspecting the individual time points' scores (Fig. S3), we found that the first component captured the major seasonal variations of climate, with scores highest during summer and lowest during winter. By contrast, the second component captured more minor seasonal variations of climate, with scores highest at the beginning of autumn and lowest at the beginning of spring. In all the analyses, we worked with these two summary components instead of the individual meteorological variables.

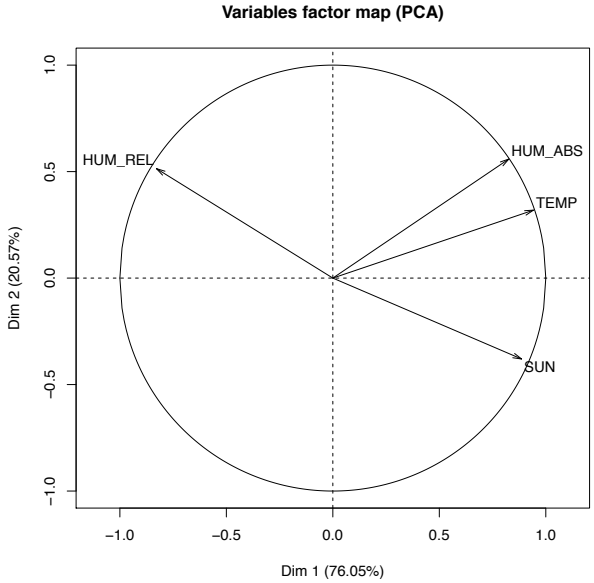

Figure S2: **Variables factor map of the PCA on meteorological data**. The meteorological variables were the temperature (TEMP), the vapor pressure (a measure of absolute humidity, HUM_ABS), the duration of insolation (SUN), and the relative humidity (HUM_REL). The loading vectors were $(0.54, 0.47, 0.51, -0.47)^{\mathrm{T}}$ for the first principal component and $(0.35, 0.62, -0.42, 0.57)^{\mathrm{T}}$ for the second principal component.

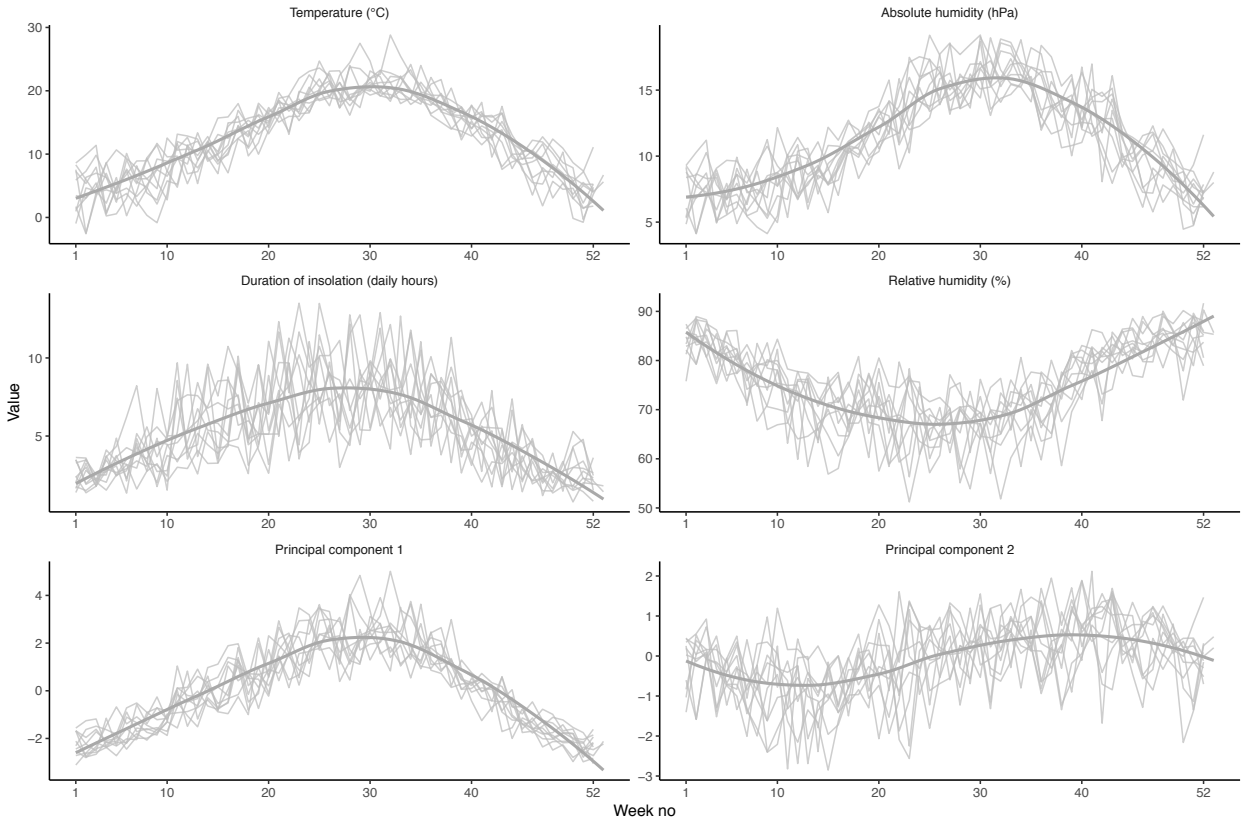

Figure S3: **Time series of the meteorological variables and the first two principal components**. In every panel, the light grey lines represent distinct years and the dark grey line the average seasonal pattern (obtained using local quadratic regression).

S2.2 Model formulation

Deterministic process model

We formulated an age-structured dynamic model of pneumococcal carriage transmission and disease. The model is an extension of the Susceptible–Carrier–Infected model and builds on models described previously [4, 5]. Although immunity to carriage was not explicitly modeled, we incorporated it indirectly by fixing age-specific durations of carriage and age-specific susceptibilities to carriage acquisition, as explained below. In every age group $i = 1, \ldots, I$, the population was divided into those individuals not carrying pneumococcus (S_i), those carrying pneumococcus, and those infected with an IPD (I_i). Because some evidence indicates that the risk of disease is higher early after carriage acquisition [6, 7], we further divided the population of carriers into those with early carriage, $C_i^{(1)}$, and late carriage, $C_i^{(2)}$. The overall duration of carriage therefore comprised a fraction of time, denoted by $f^{(1)}$, of early carriage and a fraction $1 - f^{(1)}$ of late carriage. The risk of disease was scaled by $\frac{\theta^{(1)}}{f^{(1)}}$ during early carriage and by $\frac{1-\theta^{(1)}}{1-f^{(1)}}$ during late carriage, where $\theta^{(1)}$ is the

fraction of disease that occurs during early carriage. With this parametrization, the risk of disease is higher during early carriage than during late carriage if $\frac{\theta^{(1)}}{f^{(1)}} > \frac{1-\theta^{(1)}}{1-f^{(1)}}$, or equivalently if $\theta^{(1)} > f^{(1)}$. Hence, to test the hypothesis that the disease risk is not uniform during carriage, only one parameter of the pair $(f^{(1)}, \theta^{(1)})$ needs to be estimated. We chose to fix $f^{(1)} = \frac{1}{3}$, and we estimated $\theta^{(1)}$ from the data.

To model the potential interactions between pneumococcus and ILIs, we explicitly modeled individuals infected with ILIs. We assumed that such infections occurred independently of pneumococcal carriage status, at a rate $\frac{F_i(t)}{\rho_i^{(F)}}$ in age group i. Here $F_i(t)$ is the observed weekly ILI incidence and $\rho_i^{(F)}$ the probability that a symptomatic ILI case consults a physician, assumed age-specific and fixed according to previous studies of healthcare seeking behaviors in France (Table S1 and Refs. [8, 9]). Hence, the four compartments S, $C^{(1)}$, $C^{(2)}$, and I were duplicated to represent the status of individuals vis-à-vis ILI infection: ILI-uninfected (superscript (U)) or ILI-infected (superscript (F)). According to previous evidence [10, 11, 12, 4, 13, 14], we hypothesized 3 potential mechanisms of interaction between ILIs and pneumococcus:

1. **Acquisition impact**: a non-carrier infected with an ILI was assumed to have a higher risk of carriage acquisition than an a non-carrier uninfected with ILI. The relative risk of acquisition is denoted by θ_λ.

2. **Transmission impact**: a carrier infected with an ILI was assumed to transmit at a higher rate than a carrier uninfected with an ILI. The relative risk of transmission is denoted by θ_β.

3. **Invasion impact**: a carrier infected with an ILI was assumed to have a higher risk of developing an invasive pneumococcal disease than a carrier uninfected with an ILI. The relative disease risk is denoted by θ_α.

Overall, the model consisted of 8 compartments in every age group i: $S_i^{(U)}, C_i^{(U,1)}, C_i^{(U,2)}, I_i^{(U)}, S_i^{(F)}, C_i^{(F,1)}, C_i^{(F,2)}$, and $I_i^{(F)}$. For simplicity, and because IPDs were rare during the study period, we did not explicitly model the outcome of infected individuals; rather, the infected state variables $I_i^{(U)}$ and $I_i^{(F)}$ were modeled as accumulator variables and reset to 0 at the beginning of every new week. A simplified model schematic is shown in Fig. S4.

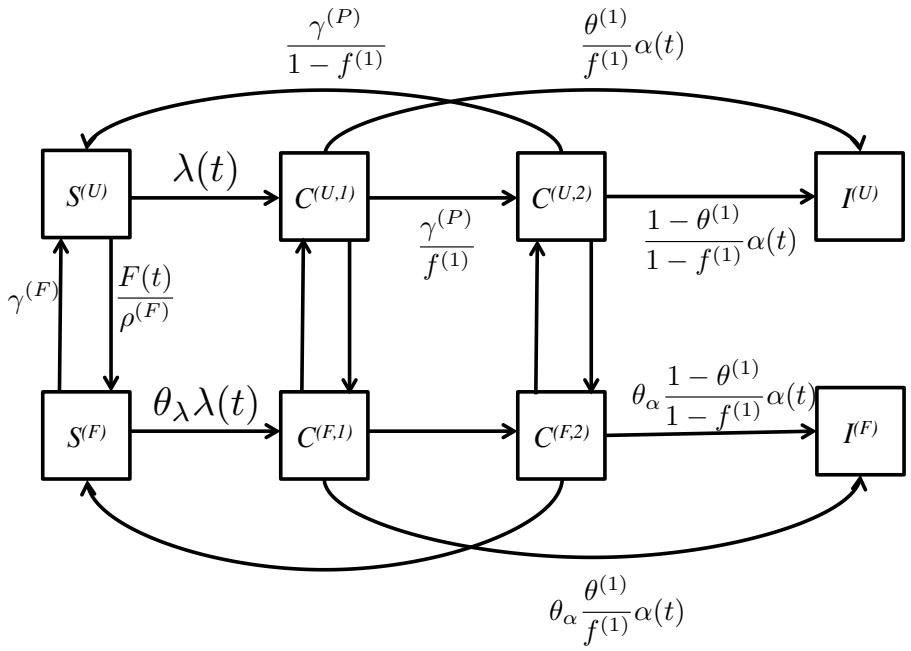

Figure S4: **Model schematic**. For simplicity, age is omitted.

The equations in the first age group (age [0,5) yr, $i = 1$) are given by:

$$
\begin{aligned}
\frac{\mathrm{d}S_1^{(U)}}{\mathrm{d}t} &= B(t) - (\lambda_1(t) + \frac{F_1(t)}{\rho_1^{(F)}})S_1^{(U)} + \frac{\gamma_1^{(P)}}{1-f^{(1)}}C_1^{(U,2)} + \gamma^{(F)}S_1^{(F)} - [\delta_1 + \mu_1(t)]S_1^{(U)} \\
\frac{\mathrm{d}C_1^{(U,1)}}{\mathrm{d}t} &= \lambda_1(t)S_1^{(U)} + \gamma^{(F)}C_1^{(F,1)} - (\frac{\gamma_1^{(P)}}{f^{(1)}} + \frac{F_1(t)}{\rho_1^{(F)}})C_1^{(U,1)} - [\delta_1 + \mu_1(t)]C_1^{(U,1)} \\
\frac{\mathrm{d}C_1^{(U,2)}}{\mathrm{d}t} &= \frac{\gamma_1^{(P)}}{f^{(1)}}C_1^{(U,1)} + \gamma^{(F)}C_1^{(F,2)} - (\frac{\gamma_1^{(P)}}{1-f^{(1)}} + \frac{F_1(t)}{\rho_1^{(F)}})C_1^{(U,2)} - [\delta_1 + \mu_1(t)]C_1^{(U,2)} \\
\frac{\mathrm{d}I_1^{(U)}}{\mathrm{d}t} &= \alpha_1(t)(\frac{\theta^{(1)}}{f^{(1)}}C_1^{(U,1)} + \frac{1-\theta^{(1)}}{1-f^{(1)}}C_1^{(U,2)}) \\
\frac{\mathrm{d}S_1^{(F)}}{\mathrm{d}t} &= \frac{F_1(t)}{\rho_1^{(F)}}S_1^{(U)} + \frac{\gamma_1^{(P)}}{1-f^{(1)}}C_1^{(F,2)} - (\theta_\lambda\lambda_1(t) + \gamma^{(F)})S_1^{(F)} - [\delta_1 + \mu_1(t)]S_1^{(F)} \\
\frac{\mathrm{d}C_1^{(F,1)}}{\mathrm{d}t} &= \theta_\lambda\lambda_1(t)S_1^{(F)} + \frac{F_1(t)}{\rho_1^{(F)}}C_1^{(U,1)} - (\gamma^{(F)} + \frac{\gamma_1^{(P)}}{f^{(1)}})C_1^{(F,1)} - [\delta_1 + \mu_1(t)]C_1^{(F,1)} \\
\frac{\mathrm{d}C_1^{(F,2)}}{\mathrm{d}t} &= \frac{\gamma_1^{(P)}}{f^{(1)}}C_1^{(F,1)} + \frac{F_1(t)}{\rho_1^{(F)}}C_1^{(U,2)} - (\gamma^{(F)} + \frac{\gamma_1^{(P)}}{1-f^{(1)}})C_1^{(F,2)} - [\delta_1 + \mu_1(t)]C_1^{(F,2)} \\
\frac{\mathrm{d}I_1^{(F)}}{\mathrm{d}t} &= \theta_\alpha\alpha_1(t)(\frac{\theta^{(1)}}{f^{(1)}}C_1^{(F,1)} + \frac{1-\theta^{(1)}}{1-f^{(1)}}C_1^{(F,2)})
\end{aligned}
$$

The equations in the other age groups ($i = 2, \ldots, I$) are:

$$
\begin{aligned}
\frac{\mathrm{d}S_i^{(U)}}{\mathrm{d}t} &= \delta_{i-1}S_{i-1}^{(U)} - (\lambda_i(t) + \frac{F_i(t)}{\rho_i^{(F)}})S_i^{(U)} + \frac{\gamma_i^{(P)}}{1-f^{(1)}}C_i^{(U,2)} + \gamma^{(F)}S_i^{(F)} - [\delta_i + \mu_i(t)]S_i^{(U)} \\
\frac{\mathrm{d}C_i^{(U,1)}}{\mathrm{d}t} &= \delta_{i-1}C_{i-1}^{(U,1)} + \lambda_i(t)S_i^{(U)} + \gamma^{(F)}C_i^{(F,1)} - (\frac{\gamma_i^{(P)}}{f^{(1)}} + \frac{F_i(t)}{\rho_i^{(F)}})C_i^{(U,1)} - [\delta_i + \mu_i(t)]C_i^{(U,1)} \\
\frac{\mathrm{d}C_i^{(U,2)}}{\mathrm{d}t} &= \delta_{i-1}C_{i-1}^{(U,2)} + \frac{\gamma_i^{(P)}}{f^{(1)}}C_i^{(U,1)} + \gamma^{(F)}C_i^{(F,2)} - (\frac{\gamma_i^{(P)}}{1-f^{(1)}} + \frac{F_i(t)}{\rho_i^{(F)}})C_i^{(U,2)} - [\delta_i + \mu_i(t)]C_i^{(U,2)} \\
\frac{\mathrm{d}I_i^{(U)}}{\mathrm{d}t} &= \alpha_i(t)(\frac{\theta^{(1)}}{f^{(1)}}C_i^{(U,1)} + \frac{1-\theta^{(1)}}{1-f^{(1)}}C_i^{(U,2)}) \\
\frac{\mathrm{d}S_i^{(F)}}{\mathrm{d}t} &= \delta_{i-1}S_{i-1}^{(F)} + \frac{F_i(t)}{\rho_i^{(F)}}S_i^{(U)} + \frac{\gamma_i^{(P)}}{1-f^{(1)}}C_i^{(F,2)} - [\theta_\lambda\lambda_i(t) + \gamma^{(F)}]S_i^{(F)} - [\delta_i + \mu_i(t)]S_i^{(F)} \\
\frac{\mathrm{d}C_i^{(F,1)}}{\mathrm{d}t} &= \delta_{i-1}C_{i-1}^{(F,1)} + \theta_\lambda\lambda_i(t)S_i^{(F)} + \frac{F_i(t)}{\rho_i^{(F)}}C_i^{(U,1)} - (\gamma^{(F)} + \frac{\gamma_i^{(P)}}{f^{(1)}})C_i^{(F,1)} - [\delta_i + \mu_i(t)]C_i^{(F,1)} \\
\frac{\mathrm{d}C_i^{(F,2)}}{\mathrm{d}t} &= \delta_{i-1}C_{i-1}^{(F,2)} + \frac{\gamma_i^{(P)}}{f^{(1)}}C_i^{(F,1)} + \frac{F_i(t)}{\rho_i^{(F)}}C_i^{(U,2)} - [\gamma^{(F)} + \frac{\gamma_i^{(P)}}{1-f^{(1)}}]C_i^{(F,2)} - [\delta_i + \mu_i(t)]C_i^{(F,2)} \\
\frac{\mathrm{d}I_i^{(F)}}{\mathrm{d}t} &= \theta_\alpha\alpha_i(t)(\frac{\theta^{(1)}}{f^{(1)}}C_i^{(F,1)} + \frac{1-\theta^{(1)}}{1-f^{(1)}}C_i^{(F,2)})
\end{aligned}
$$

Here δ_i is the aging rate in age group i, $\gamma_i^{(P)}$ the decolonization rate, $\alpha_i(t)$ the time-varying invasion rate (i.e., the rate at which carriers develop pneumococcal disease), and $\frac{1}{\gamma^{(F)}}$ the average duration of the interaction between ILIs and pneumococcus. The decolonization rates were fixed so that the average duration of carriage was 6 wk in [0,5) yr and 3 wk in older individuals [15, 16]. In keeping with previous experimental [10] and epidemiological evidence [13], we assumed a short average duration of ILI–pneumococcus interaction, fixed to $\frac{1}{\gamma^{(F)}} = \frac{2}{3}$ wk.

The carriage acquisition rate was given by

$$
\lambda_i(t) = q_i e^{\lambda_i^{(\mathrm{trend})} \times t} \sum_j c_{ij} \frac{C_j^{(U,1)} + C_j^{(U,2)} + \theta_\beta(C_j^{(F,1)} + C_j^{(F,2)})}{N_j} \mathrm{Seas}_{ij}(t)\mathrm{Christmas}_{ij}(t)
$$

where q_i is the age-specific susceptibility to carriage acquisition, $\mathrm{Seas}_{ij}(t)$ a seasonal, age-specific forcing function, and $\lambda_i^{(\mathrm{trend})}$ a trend in the acquisition rate. The latter term was used to account for strain replacement after the introduction of PCV7 and the national campaign to reduce antibiotic use in France [5]. Because the trends in IPDs differed between children aged [0,5) yr (the population targeted by PCV7) and other age groups (see also Ref. [17]), we estimated separate trend parameters in [0,5) yr and $\geq$ 5 yr. The term

$$\text{Christmas}_{ij}(t) = \begin{cases} \kappa_1(1 + A_{\text{Christmas}}) & ,\{(i = 1, \ldots, 4, j = I), (i = I, j = 1, \ldots, 4)\}, t \in T_{\text{Christmas}} \\ \kappa_1(1 - A_{\text{Christmas}}) & ,\{(i = 1, \ldots, 4, j = I), (i = I, j = 1, \ldots, 4)\}, t \notin T_{\text{Christmas}} \\ 1 & ,\text{otherwise} \end{cases}$$

represents a potential increase in contact rates between children aged [0,20) yr and the elderly during Christmas holidays, an hypothesis previously put forth to explain the early-winter peaks of IPDs [18]. The correction term $\kappa_1 = [1 - A_{\text{Christmas}}(1 - 2p_{\text{Christmas}})]^{-1}$ ($p_{\text{Christmas}} = \frac{2}{52}$, fraction of the year during Christmas holidays) is used to maintain a seasonal mean of 1. The set of time points associated with Christmas holidays, $T_{\text{Christmas}}$, was fixed according to the historical archives of school holidays in France[2]. This Christmas-related effect was incorporated into all the models, irrespective of the other hypotheses about the seasonality of carriage acquisition.

The age-specific susceptibility parameters q_i were calibrated to reach the following targets of carriage prevalence: 50% in young children [0,5) yr, 20% in children [5,20) yr, and 10% in adults and the elderly. These prevalences are broadly consistent with those observed via culture of nasopharyngeal swabs in a number of carriage studies in high-income countries [19, 20, 21, 22, 23, 24]. To perform the calibration, we considered a seasonally-unforced version of the model, simulated until endemic equilibrium. We defined a deviance function that measured the distance between the age-specific targets and the simulated prevalences of carriage. This deviance was then minimized to find the susceptibility values that best matched the target prevalences. The calibrated values are shown in Table S1 and were fixed in all subsequent model simulations.

Stochastic observation model

To complete the model specification, we modeled the observation process, which relates the model outputs to the observed data and allows to correct for under-reporting. Let $H_{i,t} = \int_{t-\Delta t}^{t} [I_i^{(U)}(t) + I_i^{(F)}(t)]dt$ represent the total simulated case count of IPDs in age group i during the time period $]t - \Delta t, t]$, where Δt is the reporting period (1 wk). The corresponding weekly case report, $D_{i,t}$, was modeled as a Negative Binomial distribution:

$$D_{i,t} \sim \text{NB}(\rho^{(P)} H_{i,t}, \frac{1}{\tau})$$

[2] http://www.education.gouv.fr/cid197/page.html

where $\rho^{(P)}$ is the average probability an IPD is reported to the surveillance system, and τ the reporting overdispersion, which adds variability to that probability. Hence, $\mathbb{E}(D_{i,t}|H_{i,t},\rho^{(P)},\tau) = \rho^{(P)}H_{i,t} = \mu$ and $\mathbb{V}(D_{i,t}|H_{i,t},\rho^{(P)},\tau) = \mu + \tau\mu^2$. The average reporting probability was fixed to $\rho^{(P)} = 0.65$, based on the results of a capture-recapture study in France [25]. This figure corresponds to an estimated coverage by the participating laboratories of 74 % of all admissions for bacterial invasive diseases in France and a near 90 % exhaustiveness of reporting of the participating laboratories.

S2.3 Tested hypotheses on seasonality

We sought to test a number of hypotheses about how seasonal variations of contacts rates, climate, and ILIs contributed to pneumococcal seasonality. Because pneumococcal disease results from carriage acquisition and subsequent invasion, we also aimed to test whether both the acquisition rate and the invasion rate varied seasonally. To do that, we first modeled a background seasonal function with annual periodicity:

$$\log s(t) = \underbrace{m_1 M_1(t-l) + m_2 M_2(t-l)}_{\text{climate-associated seasonality}} + \underbrace{\sum_{h=1}^{H}(a_h \cos \omega h t + b_h \sin \omega h t)}_{\text{unexplained seasonality}}$$

Here the left-hand side of the equation represents the seasonality associated with climate, with l the time lag (in wk) and m_1 and m_2 the meteorological coefficients. The right-hand side of the equation represents an unexplained seasonality, with $\omega = \frac{2\pi}{365/7}$ per wk the annual angular frequency, H the number of harmonics, and a_h and b_h the seasonal coefficients. This term serves as a null hypothesis, if no association with climate actually exists. We arbitrarily fixed $H = 2$ in all model simulations, but we verified that higher values did not appreciably modify our results.

We then applied the background seasonal function to different rates to test several hypotheses about the source of pneumococcal seasonality. Specifically, we tested five different models that represented different assumptions on the seasonalities of the carriage acquisition and the invasion rates:

1. **Background seasonality in carriage acquisition, no seasonality in invasion**: here we assumed no seasonality in the invasion rate and background seasonality in the carriage acquisition rate:

$$\begin{cases} \forall(i,j,t), & \text{Seas}_{ij}(t) = s(t) \\ \forall(i,t), & \alpha_i(t) = \bar{\alpha}_i \end{cases}$$

2. **No seasonality in carriage acquisition, background seasonality in invasion**: here we assumed no seasonality in the carriage acquisition rate (except for the potential increase of contacts during

Christmas) and background seasonality in the invasion rate. Hence, we wrote:

$$\begin{cases} \forall(i,j,t), & \text{Seas}_{ij}(t) = 1 \\ \forall(i,t), & \alpha_i(t) = \bar{\alpha}_i s(t) \end{cases}$$

3. **Background seasonality in acquisition, background seasonality in invasion**: here we assumed the same seasonal shape in the carriage acquisition rate and the invasion rate, but we allowed the oscillations of the acquisition rate to be damped. We hypothesized that such damped oscillations could be the result of population bottlenecks during transmission [26], if the seasonality of carriage acquisition is driven by seasonal variations in carriage density [27]. We wrote:

$$\begin{cases} \forall(i,j,t), & \log \text{Seas}_{ij}(t) = \varepsilon \log s(t) \\ \forall(i,t), & \alpha_i(t) = \bar{\alpha}_i s(t) \end{cases}$$

where the damping parameter $\varepsilon \in [0,1]$ (estimated from the data) was applied on the log-scale to preserve an annual geometric mean of 1.

4. **Term-time seasonality in acquisition, background seasonality in invasion**: here we assumed that the acquisition rate varied seasonally as a result of variations of contacts between school terms and school holidays in schoolchildren aged [0,20) yr. For simplicity, we further assumed that term-time forcing was assortative and applied only within the same age group. Hence, we wrote:

$$\text{Seas}_{ij}(t) = \begin{cases} \kappa_2(1 + A_{\text{school}}) & , i = j, i = 1, \ldots, 4, t \in T_{\text{school}} \\ \kappa_2(1 - A_{\text{school}}) & , i = j, i = 1, \ldots, 4, t \notin T_{\text{school}} \\ 1 & , \text{otherwise} \end{cases}$$

where the correction term $\kappa_2 = [1 - A_{\text{school}}(1 - 2p_{\text{school}})]$ ($p_{\text{school}} = \frac{41}{52}$, fraction of the year during school terms) was used to maintain a seasonal mean of 1. The set of times associated with school terms, T_{school}, was fixed according to historical archives of school holidays in France. The amplitude of seasonal forcing was fixed assuming that schoolchildren make 40% fewer contacts during holidays than during school terms [28], such that: $\frac{1-A_{\text{school}}}{1+A_{\text{school}}} = 0.6 \Rightarrow A_{\text{school}} = 0.25$. The age-specific invasion rate was given by $\forall(i,t), \alpha_i(t) = \bar{\alpha}_i s(t)$.

5. **Background and term-time seasonalities in acquisition, background seasonality in inva-**

sion: this model combined the hypotheses of models 3 and 4. The age-specific forcing function was therefore given by:

$$\text{Seas}_{ij}(t) = \begin{cases} s(t)^{\varepsilon} \times \kappa_2(1 + A_{\text{school}}) & ,i = j, i = 1, \ldots, 4, t \in T_{\text{school}} \\ s(t)^{\varepsilon} \times \kappa_2(1 - A_{\text{school}}) & ,i = j, i = 1, \ldots, 4, t \notin T_{\text{school}} \\ s(t)^{\varepsilon} & ,\text{otherwise} \end{cases}$$

and the age-specific invasion rate by: $\forall(i,t), \alpha_i(t) = \bar{\alpha}_i s(t)$.

The list of fixed model parameters is presented in Table S1.

Parameter	Meaning	Value	Source
q_i	Susceptibility to carriage acquisition	$q_1 = 0.015$, $q_{2,3,4} = 0.004$, $q_{5,6,7,8} = 0.003$, $q_9 = 0.005$	Calibrated to reach a target carriage prevalence in each age group
$1/\gamma_i^{(P)}$	Carriage durations	6 wk in [0, 5) yr, 3 wk in ≥ 5 yr	[16, 15]
δ_i	Aging rates	$\delta_{i=1,\ldots,4} = \frac{1}{5}\,\text{yr}^{-1}$, $\delta_{i=5,\ldots,8} = \frac{1}{10}\,\text{yr}^{-1}$, $\delta_9 = \frac{1}{20}\,\text{yr}^{-1}$	—
C_i^*	Target prevalences	0.5 in [0,5) yr, 0.2 in [5,20) yr, 0.1 in ≥ 20 yr	[29, 30]
c_{ij}	Age-specific contact rates	Fig. S1	[1]
$1/\gamma^{(F)}$	Duration of interaction	$\frac{2}{3}$ wk	Assume short-lived interaction [10, 13]
$\rho_i^{(F)}$	Fraction of ILI cases that consult a physician	1 in [0,5) yr, 0.45 in [15,30) yr, 0.6 in other age groups	[9, 8]
$\rho^{(P)}$	Probability that an IPD case is reported	0.65	[25]
$f^{(1)}$	Fraction of carriage duration in high-risk groups	1/3	Can be fixed to an arbitrary value, as long as $\theta^{(1)}$ is estimated
A_{school}	Amplitude of term-time forcing	0.25	[28]
H	No of harmonics to model the function representing unknown seasonality	2	Assumption

Table S1: **Fixed model parameters**.

S2.4 Estimation procedure

For every hypothesis regarding the seasonality of carriage acquisition, we conducted maximum likelihood estimation using trajectory matching [31]. This was done in 4 steps:

1. We first estimated the model without association with climate and ILIs. We ran 2,500 replicates estimations, initiated from 2,500 starting parameter sets generated using Latin hypercube sampling over a broad range of parameters.

2. We tested the models without ILIs, but with climate incorporated at a given time lag, $l = 0, \ldots, 3$ wk. For each time lag, we ran 2,500 replicate estimations from 2,500 starting parameter sets. These starting parameters sets were generated using Latin hypercube sampling over the multivariate 95% confidence interval of the parameters estimated at the previous step.

3. We identified the best model from step 2 and included ILIs, with all the interaction parameters (i.e., θ_α, θ_β, and θ_λ) estimated simultaneously. We ran 2,500 replicate estimations from 2,500 starting parameter sets, whose ranges were narrowed using the parameters estimated at the previous step.

4. If one or more mechanisms of interaction were identified in the previous step, the corresponding interaction parameters were made age-specific. In that case, for each interaction mechanism identified we estimated three interaction parameters in individuals aged [0,5), [5,60), and 60+ yr, denoted by $\theta_{.,1}$, $\theta_{.,\{2,\ldots,8\}}$, and $\theta_{.,9}$.

For every model, the estimations were repeated from the best 10% parameter sets to pinpoint the maximum likelihood estimate. The final convergence was assessed by inspecting the sliced log-likelihood around every estimated parameter. Finally, we used a parametric bootstrap to assess parametric uncertainty and to calculate approximate 95% confidence intervals around the maximum likelihood estimates [32].

S2.5 Simulation protocol and implementation details

All model simulations were started in 1985, assuming that the system was at equilibrium before the first data point (wk 27 or year 2000). The set of differential equations was integrated numerically; the log-likelihood was maximized using the "Subplex" algorithm, implemented in the R nloptr package [33]. All the analyses were performed using the pomp package (version 1.12) [31], which operates in the R environment [34]. The R checkpoint package [35] was used to freeze the packages' version and to ensure reproducibility of the results.

S3 Supplementary results

S3.1 Models comparison and best model's parameter estimates

The AIC, log-likelihood, and R^2 of all the models tested are presented in Table S2. The parameter estimates of the best model (i.e., with lowest AIC) are shown in Table S3.

Hypothesis about source of seasonality		**Seasonal covariates included**								
Acquisition rate	Invasion rate	Climate	Ø	Lag 0	Lag –1	Lag –2	Lag –3	Lag –1	Lag –1	Lag –1
		ILIs	Ø	Ø	Ø	Ø	Ø	Lag –1	Lag 0	Lag 0
		ILI Interaction parameters	Ø	Ø	Ø	Ø	Ø	All	All	All, age-specific θ_α
Background seasonality	Ø	AIC	15252.1	15078.1	15112.7	15175.5	15207.5	—	—	—
		log L	–7612.1	–7523.1	–7540.4	–7571.8	–7592.7			
		n_θ	14	16	16	16	16			
		R^2	0.72	0.76	0.75	0.74	0.73			
Ø	Background seasonality	AIC	15211.2	15162.1	15079.9	15097.3	15172.9	14965.9	14917.5	**14890.2**
		log L	–7591.6	–7565.1	–7524.0	–7532.7	–7570.5	–7463.9	–7439.7	–7424.1
		n_θ	14	16	16	16	16	19	19	21
		R^2	0.72	0.74	0.76	0.75	0.73	0.76	0.77	0.78
Background seasonality	Background seasonality	AIC	15180.4	15079.4	15011.8	15049.9	15133.4	14900.2	14854.6	**14837.2**
		log L	–7575.2	–7522.7	–7488.9	–7508.0	–7549.7	–7430.1	–7407.3	–7396.6
		n_θ	15	17	17	17	17	20	20	22
		R^2	0.73	0.75	0.77	0.75	0.74	0.78	0.78	0.78
Term-time seasonality	Background seasonality	AIC	15145.5	15096.3	15010.0	15027.7	15105.5	14906.3	14858.6	**14826.6**
		log L	–7558.7	–7532.1	–7489.0	–7497.8	–7536.7	–7434.1	–7410.3	–7392.3
		n_θ	14	16	16	16	16	19	19	21
		R^2	0.73	0.74	0.76	0.75	0.74	0.77	0.78	0.78
Background and term-time seasonalities	Background seasonality	AIC	**15115.8**	**14996.8**	**14935.8**	**14981.1**	**15066.1**	**14835.3**	**14790.9**	**14773.6**
		log L	–7542.9	–7481.4	–7450.9	–7473.6	–7516.1	–7397.7	–7375.4	–7364.8
		n_θ	15	17	17	17	17	20	20	22
		R^2	0.73	0.76	0.77	0.76	0.74	0.78	0.79	0.79

Table S2: **Models comparison**. The table shows the AIC, log-likelihood, number of estimated parameters n_θ, and R^2 of models representing different assumptions on the mechanisms underlying IPD seasonality (i.e., seasonality in the invasion rate or in both the invasion rate and the carriage acquisition rate) and the association with climate and ILIs, incorporated as covariates in the corresponding models. The background seasonality refers to a time-varying function with annual periodicity, which was modeled as a Fourier function with annual and semi-annual harmonics and which additionally incorporated the climate variables shown in Fig. 1C, depending on the model tested. The models with lowest AIC are indicated in boldface, for every line and every column of the table.

Parameter	Meaning	Estimate (95% CI)
$\bar{\alpha}_1$	Average invasion rate in [0,5) yr	6.8 (6.3, 7.3)$\times 10^{-6}$ wk^{-1}
$\bar{\alpha}_{2,3,4}$	Average invasion rate in [5,20) yr	1.9 (1.8, 2.0)$\times 10^{-6}$ wk^{-1}
$\bar{\alpha}_{5,6}$	Average invasion rate in [20,40) yr	5.9 (5.7, 6.2)$\times 10^{-6}$ wk^{-1}
$\bar{\alpha}_{7,8}$	Average invasion rate in [40,60) yr	1.8 (1.7, 1.8)$\times 10^{-5}$ wk^{-1}
$\bar{\alpha}_9$	Average invasion rate in 60+ yr	4.8 (4.6, 4.9)$\times 10^{-5}$ wk^{-1}
$\lambda_1^{(\text{trend})}$	Trend in acquisition rate in [0,5) yr (effect of PCV7)	–3.6 (–6.2, –2.1)$\times 10^{-4}$ wk^{-1}
$\lambda_{2,\ldots,9}^{(\text{trend})}$	Trend in acquisition rate in 5+ yr (effect of PCV7)	1.1 (0.8, 1.4)$\times 10^{-4}$ wk^{-1}
m_1	Meteorological coefficients	–0.05 (–0.08, –0.03)
m_2		–0.12 (–0.15, –0.10)
a_1	Coefficients of first harmonic	–0.15 (–0.212, –0.10)
b_1		0.06 (0.00, 0.09)
a_2	Coefficients of second harmonic	–0.06 (–0.08, 0.04)
b_2		–0.09 (–0.12, –0.06)
ε	Damping coefficient of carriage acquisition rate	0.30 (0.13, 0.43)
$A_{\text{Christmas}}$	Amplitude of contact rates increase during Christmas	0.74 (0.69, 0.79)
$\theta^{(1)}$	Fraction of disease during early carriage	1.00 (0.72, 1.00)
τ	Reporting overdispersion	0.029 (0.024, 0.032)
$\theta_{\alpha,1}$	Relative invasion risk in ILI-infected aged [0,5)	49 (30, 68)
$\theta_{\alpha,\{2,\ldots,8\}}$	Relative invasion risk in ILI-infected aged [5,60)	59 (40, 72)
$\theta_{\alpha,9}$	Relative invasion risk in ILI-infected aged 60+	146 (89, 188)
θ_β	Relative transmission risk in ILI-infected of any age	1.0 (1.0, 3.3)
θ_λ	Relative acquisition risk in ILI-infected of any age	1.0 (1.0, 2.1)

Table S3: **Best model parameter estimates**. The confidence intervals are based on 144 bootstrap replicates.

S3.2 Model assessment

Model predictions in France

To assess the best model fit, we first compared model simulations with the observed data. As shown in Fig. S5A, this comparison indicated excellent model-data agreement. Inspecting the model prediction residuals (Fig. S5B), we note, however, that the best model still tended to under-estimate IPD peaks around week no 40 in young children. To further quantify model-data agreement, we calculated a generalized R^2 [36], defined as:

$$R^2 = 1 - \frac{\sum_{i,t}(D_{i,t} - \rho^{(P)} H_{i,t})^2}{\sum_{i,t}(D_{i,t} - \bar{D}_i)^2}$$

where $D_{i,t}$ is the observed number of IPDs in age group i during wk t and $\bar{D}_i = \frac{1}{T}\sum_t D_{i,t}$ the time-averaged observed number of IPDs in age group i. The R^2 of the best model equalled 0.79; the R^2 of all the other models are also presented in Table S2.

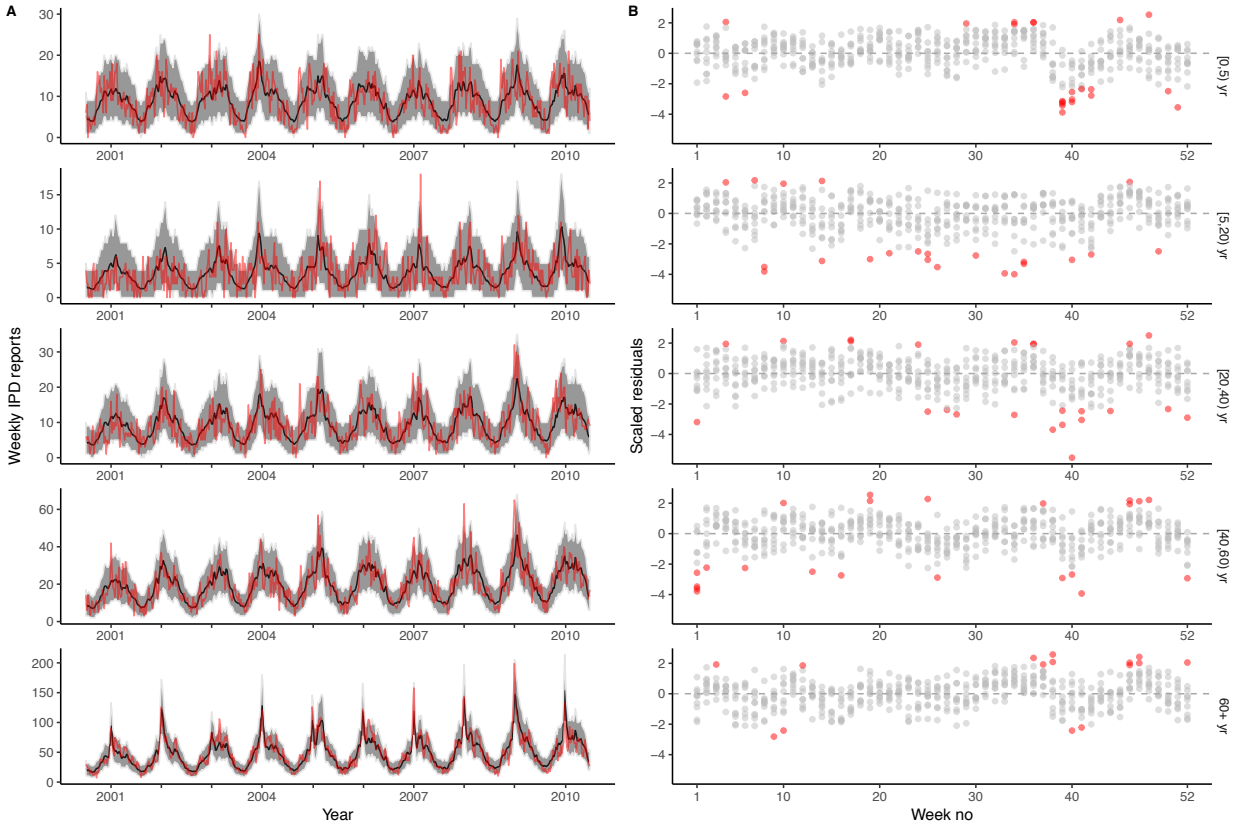

Figure S5: **Best model fit to data**. (A) For every age group, the average deterministic skeleton (black line, corrected for under-reporting) and the prediction envelope (grey ribbon, based on the 2.5% and 97.5 quantiles of the negative-binomial observation model) are displayed. The red lines show the observed data. For visual clarity, the y-axis values differ between panels. (B) Scaled residuals in every age group, according to week no. The red dots indicate data points that were outside of the prediction envelope.

Model predictions in French regions

We examined the model fit on region-level IPD incidence data in France. To do this, we constructed tables of covariates (i.e., birth rate, ILIs, and climate) in five geographical regions spanning mainland France, as defined previously [37]: Ile-de-France (including Paris), Northwest, Northeast, Southeast, and Southwest. We then simulated all the models (without re-estimating the parameters) and calculated the log-likelihood and the R^2 as before. Encouragingly, we found our results to be robust in every region, with highest log-likelihood for the model that incorporated ILIs and climate (Table S4). Although no parameters were estimated, the model's predictive power remained acceptable, with R^2 ranging from 0.39 in the Ile-de-France region to 0.56 in the Northwest region.

Region	Model tested								
	Climate	Ø	Lag 0	Lag –1	Lag –2	Lag –3	Lag –1	Lag –1	Lag –1
	ILIs	Ø	Ø	Ø	Ø	Ø	Lag –1	Lag 0	Lag 0
	Interaction parameters	Ø	Ø	Ø	Ø	Ø	All	All	All, age-specific θ_α
Ile-de-France	D	9878.0	9890.7	9875.5	9883.0	9889.9	9840.5	9828.6	**9823.3**
	R^2	0.36	0.36	0.37	0.36	0.36	0.38	0.39	0.39
Northwest	D	9598.8	9543.8	9518.6	9528.0	9563.1	9477.0	**9468.4**	9473.9
	R^2	0.51	0.54	0.55	0.54	0.52	0.56	0.56	0.56
Northeast	D	10215.8	10164.4	10168.2	10186.2	10213.8	10149.5	10138.8	**10118.9**
	R^2	0.51	0.54	0.53	0.52	0.51	0.54	0.54	0.55
Southeast	D	10231.4	10205.1	10196.7	10205.4	10220.4	10165.3	10134.3	**10121.9**
	R^2	0.49	0.52	0.52	0.51	0.50	0.52	0.54	0.54
Southwest	D	8210.2	8216.6	8198.0	8208.3	8218.3	8180.3	8165.5	**8161.9**
	R^2	0.48	0.48	0.49	0.48	0.48	0.50	0.50	0.50

Table S4: **Model fit on region-level IPD incidence data**. For every region, the table shows the values of the R^2 and of the deviance, defined as $D = -2 \log L$. The model with lowest deviance is indicated in boldface for every region. These fits are for the models with background and term-time seasonalities in carriage acquisition; the association with climate and ILIs was also robust for the other hypotheses regarding the seasonality of acquisition.

Model predictions in tropical climates

We sought to (qualitatively) compare our model predictions with data from two recent studies that examined the seasonality of pneumococcal carriage in young children, in Mae Sot, Thailand [38] and in Sibanor, the Gambia [39]. Both locations have a tropical climate (Köppen–Geiger group Aw), characterized by a hot dry season and a rainy season (from May to October in Mae Sot, and from June to October in Sibanor, cf. Fig. S6). In both studies, the prevalence of pneumococcal carriage was found to be higher during the dry season than during the rainy season. To determine if our model was able to reproduce these variations, we proceeded in several steps. First, we collected weekly records of the average temperature and the average relative humidity from weather stations near each location (Tambacounda airport, 136 km east of Sibanor; and Chiang Mai International Airport, 232 km north of Mae Sot) during 2014–2017[3]. Because the other climatic variables used in our analysis (vapor pressure and duration of sunshine) were unavailable, we re-fitted the models in France to get parameter estimates for the best model, based on a PCA with temperature and relative humidity only. (We note, however, that the principal components derived from that PCA were almost identical [correlation coefficients exceeding 95%] to those from the original PCA.) For these estimations, we also re-calibrated the susceptibility parameters so that the average prevalences of carriages were 65% in [0,5), 45% in [5,20), and 30% in $\geq$ 20 yr, based on estimates in low-income countries [40, 41].

We then ran model simulations during 2014–2017 using the PCA-transformed climatic data in Mae Sot or Sibanor and the climatic parameters re-estimated in France. In the absence of further data, we deactivated all the other seasonal factors (that is, ILIs, seasonal variations of contact rates, and the unexplained seasonality)

[3] The data were downloaded from WolframAlpha (https://www.wolframalpha.com/) using the queries 'temperature Sibanor' and 'temperature Mae Sot'.

and we used the French contact matrix. These caveats in mind, the resulting model predictions of carriage prevalence in children aged [0,5) yr are shown in Fig. S6. In qualitatively good agreement with the data in both locations [39, 38], the model predicted smooth variations of carriage prevalence over the year, with a peak in January-February during the dry season and a trough in June–July during the rainy season. Although a more quantitative comparison is made difficult by the lack of data about the other seasonal factors, we note that the seasonal amplitude predicted by our model (10–15% difference of prevalence between the dry and the rainy seasons) is roughly consistent with that estimated in Refs. [39, 38] (e.g., seasonal amplitude of ≈17% for carriage of any serotype in children 2–4 yr, Fig. 2 in Ref. [39]).

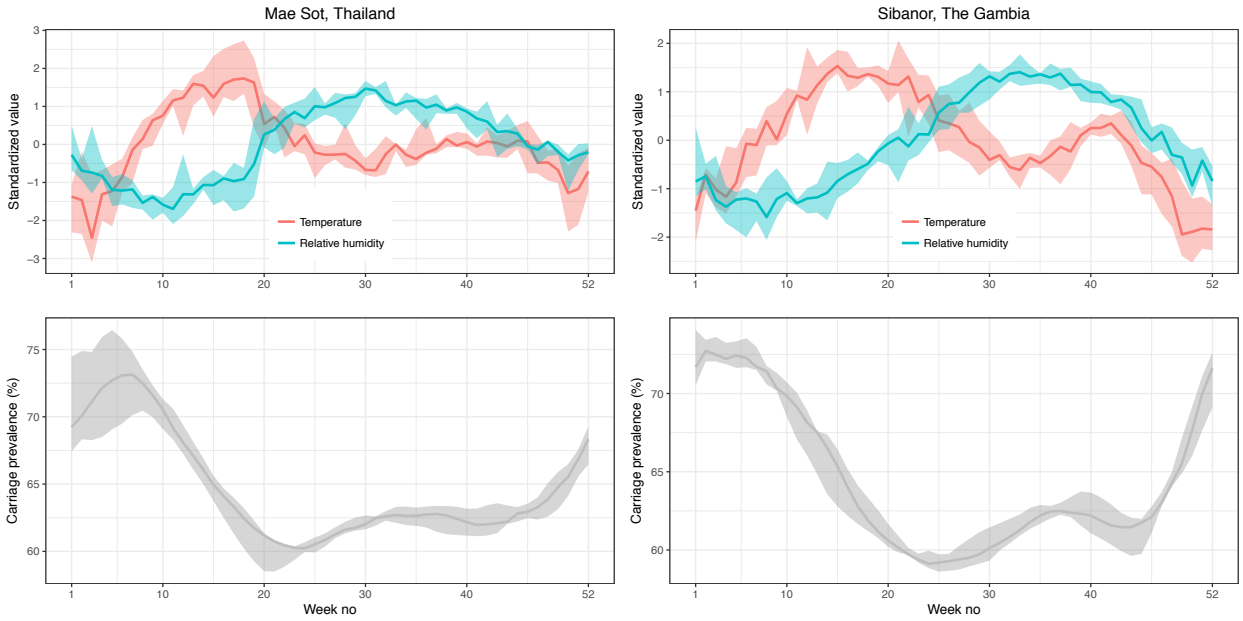

Figure S6: Model out-of-fit predictions in Mae Sot, Thailand [38] and Sibanor, The Gambia [39]. Top row: Standardized values of average temperature and average relative humidity. Bottom: simulated carriage prevalence in children <5 yr. In each panel, the line (envelope) represents the median (range) values during 2014–2017.

S3.3 Sensitivity analyses

Different contact matrix

To assess the robustness of our results, we re-estimated all the models with a contact matrix derived from the POLYMOD study in Great Britain, displayed in Fig. S7. The calibrated susceptibilities to carriage acquisition with this contact matrix were 0.014 in [0,5) yr, 0.004 in [5,20) yr, 0.003 in [20,60) yr, and 0.006 in 60+ yr. As shown in Table S5, the results based on this contact matrix were almost similar to those of the base analysis.

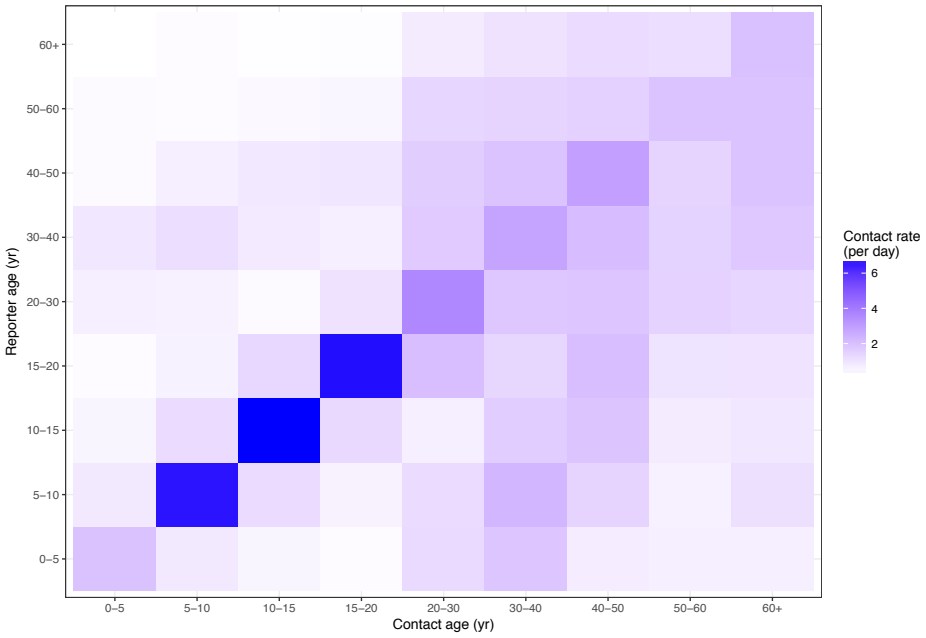

Figure S7: **Matrix of age-specific contact rates in Great Britain**. Data are from Ref. [2].

Hypothesis about source of seasonality		**Seasonal covariates included**							
Acquisition rate	Invasion rate	Climate	Ø	Lag 0	Lag –1	Lag –2	Lag –3	Lag –1	Lag –1
		ILIs	Ø	Ø	Ø	Ø	Ø	Lag 0	Lag 0
		ILI Interaction parameters	Ø	Ø	Ø	Ø	Ø	All	All, age-specific θ_α
Ø	Background seasonality	AIC	15211.4	15162.7	15080.3	15096.4	15173.2	14920.3	**14892.3**
		log L	–7591.7	–7565.3	–7524.1	–7532.2	–7570.6	–7441.1	–7425.2
		n_θ	14	16	16	16	16	19	21
Background seasonality	Background seasonality	AIC	15180.7	15078.7	15010.7	15051.5	15133.9	14858.3	**14839.6**
		log L	–7575.4	–7522.4	–7488.4	–7508.8	–7549.9	–7409.1	–7397.8
		n_θ	15	17	17	17	17	20	22
Term-time seasonality	Background seasonality	AIC	15136.7	15088.0	15001.7	15017.4	15096.5	14854.1	**14821.3**
		log L	–7554.4	–7528.0	–7484.9	–7492.7	–7532.3	–7408.1	–7389.6
		n_θ	14	16	16	16	16	19	21
Background and term-time seasonalities	Background seasonality	AIC	**15113.2**	**15004.5**	**14937.9**	**14978.7**	**15064.9**	**14793.5**	**14774.5**
		log L	–7541.6	–7485.3	–7451.9	–7472.3	–7515.5	–7376.8	–7365.3
		n_θ	15	17	17	17	17	20	22

Table S5: **Models comparison (contact matrix from Great Britain)**. The models with lowest AIC are indicated in boldface, for every line and every column of the table.

Different reporting probability in the elderly

To verify the robustness of our results regarding the age-specific interaction of ILIs with pneumococcus (in particular in the elderly), we re-fitted the best model assuming a reporting probability of 50% or 80% in the elderly and of 65% in the other age groups. As shown in Table S6, the variations in reporting probability were compensated by a change in the average invasion rate in the elderly (parameter $\bar{\alpha}_9$), but the age-specific interaction parameters changed little, if at all. Hence, our main results were robust to these alternative

assumptions.

Model tested	$\log L$	θ_β	θ_λ	$\theta_{\alpha,1}$	$\theta_{\alpha,\{2,\ldots,8\}}$	$\theta_{\alpha,9}$	$\bar{\alpha}_9$ ($\times 10^{-5}$ wk^{-1})
Reporting probability of 65% in the elderly (base model)	–7364.8	1	1	49	59	146	4.8
Reporting probability of 50% in the elderly (base model)	–7364.8	1	1	50	58	144	6.0
Reporting probability of 80% in the elderly (base model)	–7364.9	1	1	50	58	143	3.7

Table S6: **Point parameter estimates for alternative assumptions regarding the reporting probability in the elderly**.

Interaction estimated with bronchiolitis data in children <5 yr

In addition to influenza, other respiratory viruses—such as the respiratory syncytial virus (RSV)—may interact with pneumococcus [42]. To examine this, we gathered incidence data on bronchiolitis (ICD-10 code J21, here considered as a proxy for RSV infection) in children <5 yr. These data were available from the Oscour network, a surveillance network set up after the 2003 heat wave in France and based on admissions to hospital emergency departments [43]. The number of hospital emergency departments participating in the network increased gradually over time, with an exhaustiveness estimated at 14% in 2007, 18% in 2008, 25% in 2009, and 45% in 2010 (unpublished estimates). We here considered the data during years 2007/08–2009/10, corrected for under-reporting using those estimates of exhaustiveness (Fig. S8).

We then fitted the interaction model as before. Because of the limited amount of data, however, we tested each interaction mechanism separately. The results, presented in Table , suggested the existence of an interaction qualitatively and quantitatively comparable to that with ILIs (interaction parameter point estimate of 78 for the invasion impact hypothesis; of 1 for the other hypotheses). Nevertheless, the difference with the no-interaction model ($\log L = -7450.9$, cf. Table S2) was not statistically significant ($\log L = -7450.6$, $\Delta \log L = 0.3$, likelihood ratio test $P = 0.44$), presumably as a result of lack of statistical power.

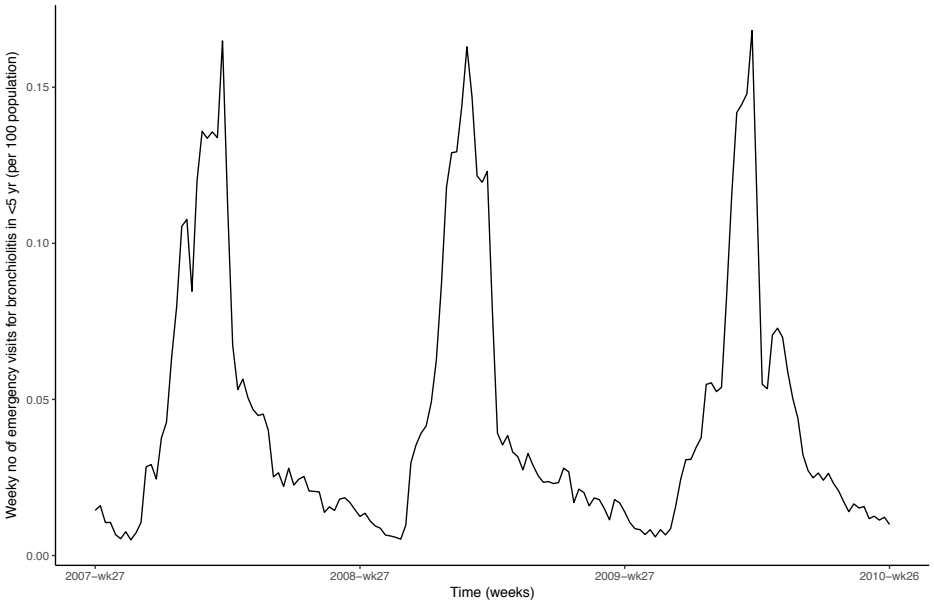

Figure S8: **Weekly incidence of hospital emergency department visits for bronchiolitis (ICD-10 code J21) in children <5 yr**. The data were corrected for under-reporting, as explained above.

Model	$\log L$	Interaction parameter point estimate
No-interaction	–7450.9	—
Invasion impact	–7450.6	78
Acquisition impact	–7450.9	1
Transmission impact	–7450.9	1

Table S7: **Interaction parameter point estimates with bronchiolitis data in children <5 yr**.

S3.4 Simulation protocol of the seasonally-timed intervention

We simulated a seasonally-timed intervention by reducing the contact rates involving a target age group ([0,5), [5,20), [20,40), [40,60), or 60+ yr) during a target week no (between 1 and 52) throughout the study period. Specifically, we assumed a control effort that consisted of reducing the total number of daily contacts by 1% during every day of the target week. Let E represent the total number of daily contacts prevented by that control effort and E_i the total number of daily contacts involving the age group i targeted by the intervention. Because E_i varied with age, we maintained a constant control effort by reducing the contact rates of the target age group with all the age groups by a factor $\frac{E}{E_i}$. Therefore, the contact rates of the target age group were modified such that: $\forall j, c_{ij} \leftarrow (1 - \frac{E}{E_i})c_{ij}$. The symmetrical contact rates c_{ji} were similarly decreased to ensure the reciprocity of contacts. The model was then simulated throughout the study period. The resulting age-specific numbers of IPDs were calculated and compared to those of the base scenario, without the seasonally-timed intervention. The results of these experiments are presented in

Fig. S9; the main text Fig. 3C is a subset of this figure, focused on the young children and the elderly. The results based on the best model's estimates with the POLYMOD contact matrix in Great Britain (Fig. S7 and Table S5) are shown in Fig. S10.

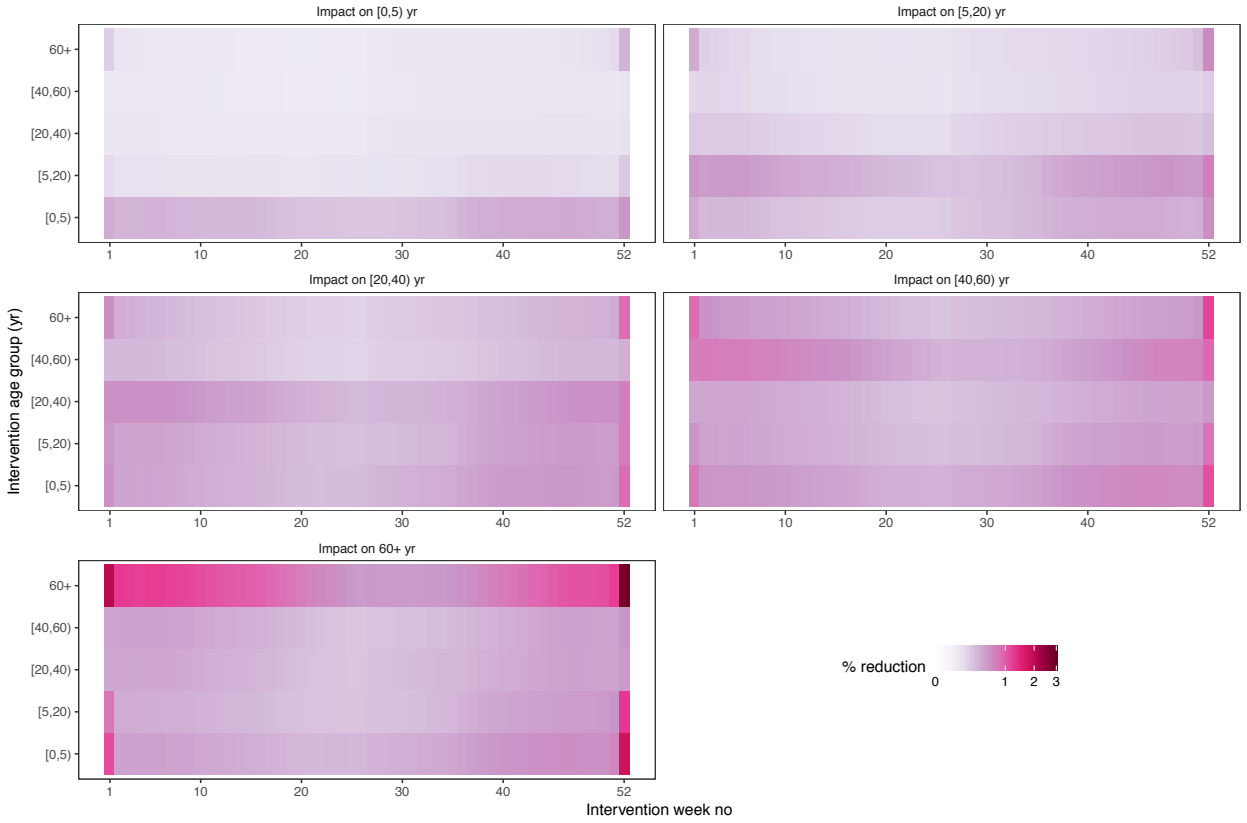

Figure S9: **Predicted impact of a seasonally-timed intervention (contact matrix from France)**. The figure shows the predicted relative reduction in IPDs caused by a seasonally-timed intervention during a given week no (x-axis) in a given target age group (y-axis). The different panels show the predicted impact of the intervention in different age groups. For visual clarity, the color scale is square-root transformed.

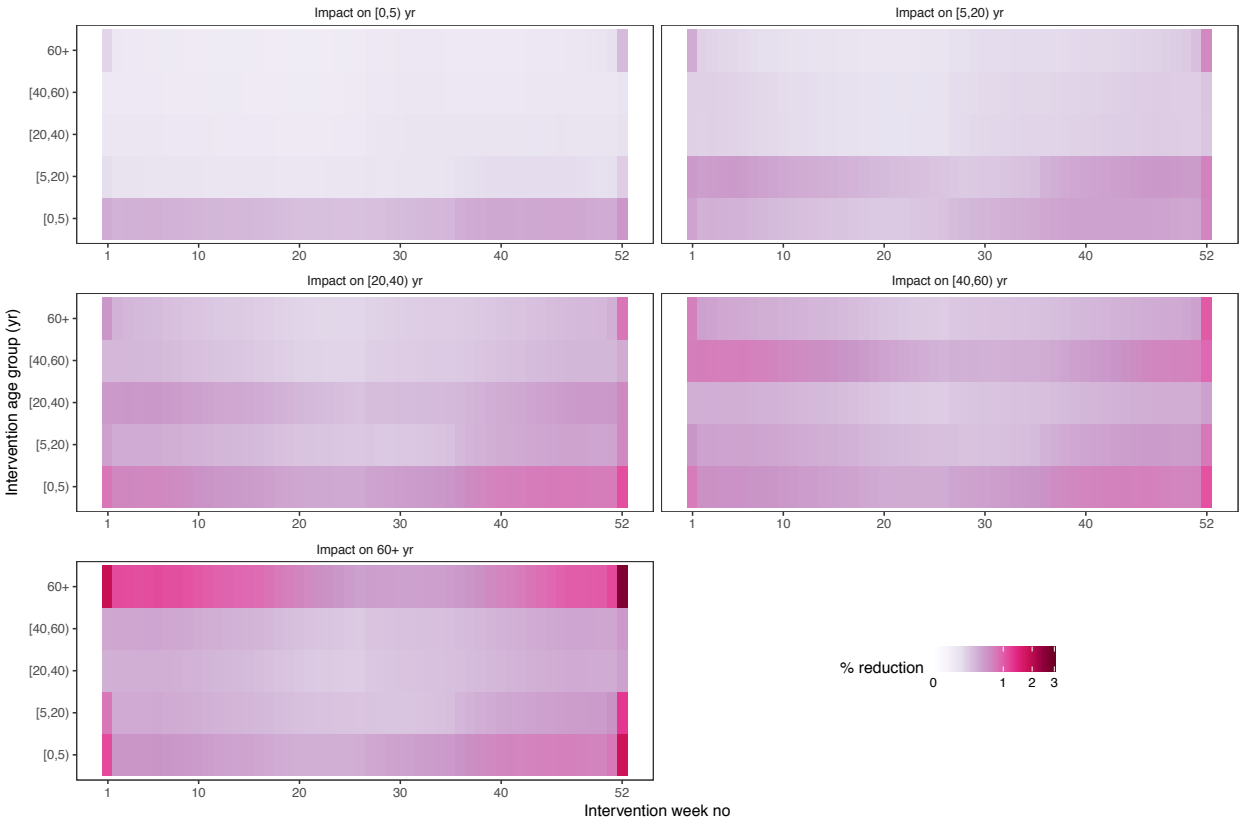

Figure S10: **Predicted impact of a seasonally-timed intervention (contact matrix from Great Britain)**. The figure shows the predicted relative reduction in IPDs caused by a seasonally-timed intervention during a given week no (x-axis) in a given target age group (y-axis). The different panels show the predicted impact of the intervention in different age groups. For visual clarity, the color scale if square-root transformed.

References

[1] Béraud, G, Kazmercziak, S, Beutels, P, Levy-Bruhl, D, Lenne, X, Mielcarek, N, Yazdanpanah, Y, Boëlle, P.-Y, Hens, N, & Dervaux, B. (2015) The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLoS One* **10**, e0133203.

[2] Mossong, J, Hens, N, Jit, M, Beutels, P, Auranen, K, Mikolajczyk, R, Massari, M, Salmaso, S, Tomba, G. S, Wallinga, J, Heijne, J, Sadkowska-Todys, M, Rosinska, M, & Edmunds, W. J. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* **5**, e74.

[3] Lê, S, Josse, J, & Husson, F. (2008) FactoMineR: A package for multivariate analysis. *Journal of Statistical Software* **25**, 1–18.

[4] Opatowski, L, Varon, E, Dupont, C, Temime, L, van der Werf, S, Gutmann, L, Boëlle, P.-Y, Watier, L, & Guillemot, D. (2013) Assessing pneumococcal meningitis association with viral respiratory infections and antibiotics: insights from statistical and mathematical models. *Proc Biol Sci* **280**, 20130519.

[5] Domenech de Cellès, M, Pons-Salort, M, Varon, E, Vibet, M.-A, Ligier, C, Letort, V, Opatowski, L, & Guillemot, D. (2015) Interaction of vaccination and reduction of antibiotic use drives unexpected increase of pneumococcal meningitis. *Sci Rep* **5**, 11293.

[6] Gray, B. M, Converse, 3rd, G. M, & Dillon, Jr, H. C. (1980) Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. *J Infect Dis* **142**, 923–33.

[7] Syrjänen, R. K, Auranen, K. J, Leino, T. M, Kilpi, T. M, & Mäkelä, P. H. (2005) Pneumococcal acute otitis media in relation to pneumococcal nasopharyngeal carriage. *Pediatr Infect Dis J* **24**, 801–6.

[8] Carrat, F, Sahler, C, Rogez, S, Leruez-Ville, M, Freymuth, F, Le Gales, C, Bungener, M, Housset, B, Nicolas, M, & Rouzioux, C. (2002) Influenza burden of illness: estimates from a national prospective survey of household contacts in France. *Arch Intern Med* **162**, 1842–8.

[9] Van Cauteren, D, Vaux, S, de Valk, H, Le Strat, Y, Vaillant, V, & Lévy-Bruhl, D. (2012) Burden of influenza, healthcare seeking behaviour and hygiene measures during the A(H1N1)2009 pandemic in France: a population based study. *BMC Public Health* **12**, 947.

[10] McCullers, J. A & Rehg, J. E. (2002) Lethal synergism between influenza virus and *Streptococcus pneumoniae*: characterization of a mouse model and the role of platelet-activating factor receptor. *J Infect Dis* **186**, 341–50.

[11] McCullers, J. A, McAuley, J. L, Browall, S, Iverson, A. R, Boyd, K. L, & Henriques Normark, B. (2010) Influenza enhances susceptibility to natural acquisition of and disease due to *Streptococcus pneumoniae* in ferrets. *J Infect Dis* **202**, 1287–95.

[12] Diavatopoulos, D. A, Short, K. R, Price, J. T, Wilksch, J. J, Brown, L. E, Briles, D. E, Strugnell, R. A, & Wijburg, O. L. (2010) Influenza a virus facilitates *Streptococcus pneumoniae* transmission and disease. *FASEB J* **24**, 1789–98.

[13] Shrestha, S, Foxman, B, Weinberger, D. M, Steiner, C, Viboud, C, & Rohani, P. (2013) Identifying the interaction between influenza and pneumococcal pneumonia using incidence data. *Sci Transl Med* **5**, 191ra84.

[14] Shrestha, S, Foxman, B, Berus, J, van Panhuis, W. G, Steiner, C, Viboud, C, & Rohani, P. (2015) The role of influenza in the epidemiology of pneumonia. *Sci Rep* **5**, 15314.

[15] Högberg, L, Geli, P, Ringberg, H, Melander, E, Lipsitch, M, & Ekdahl, K. (2007) Age- and serogroup-related differences in observed durations of nasopharyngeal carriage of penicillin-resistant pneumococci. *J Clin Microbiol* **45**, 948–52.

[16] Melegaro, A, Gay, N. J, & Medley, G. F. (2004) Estimating the transmission parameters of pneumococcal carriage in households. *Epidemiol Infect* **132**, 433–41.

[17] Alari, A, Chaussade, H, Domenech De Cellès, M, Le Fouler, L, Varon, E, Opatowski, L, Guillemot, D, & Watier, L. (2016) Impact of pneumococcal conjugate vaccines on pneumococcal meningitis cases in france between 2001 and 2014: a time series analysis. *BMC Med* **14**, 211.

[18] Walter, N. D, Taylor, Jr, T. H, Dowell, S. F, Mathis, S, Moore, M. R, & Active Bacterial Core Surveillance System Team. (2009) Holiday spikes in pneumococcal disease among older adults. *N Engl J Med* **361**, 2584–5.

[19] Henriqus Normark, B, Christensson, B, Sandgren, A, Noreen, B, Sylvan, S, Burman, L. G, & Olsson-Liljequist, B. (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**, 337–44.

[20] Hussain, M, Melegaro, A, Pebody, R. G, George, R, Edmunds, W. J, Talukdar, R, Martin, S. A, Efstratiou, A, & Miller, E. (2005) A longitudinal household study of *Streptococcus pneumoniae* nasopharyngeal carriage in a UK setting. *Epidemiol Infect* **133**, 891–8.

[21] Kaltoft, M. S, Skov Sørensen, U. B, Slotved, H.-C, & Konradsen, H. B. (2008) An easy method for detection of nasopharyngeal carriage of multiple *Streptococcus pneumoniae* serotypes. *J Microbiol Methods* **75**, 540–4.

[22] Greenberg, D, Broides, A, Blancovich, I, Peled, N, Givon-Lavi, N, & Dagan, R. (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**, 4604–9.

[23] Regev-Yochay, G, Raz, M, Dagan, R, Porat, N, Shainberg, B, Pinco, E, Keller, N, & Rubinstein, E. (2004) Nasopharyngeal carriage of *Streptococcus pneumoniae* by adults and children in community and family settings. *Clin Infect Dis* **38**, 632–9.

[24] Palmu, A. A, Kaijalainen, T, Saukkoriipi, A, Leinonen, M, & Kilpi, T. M. (2012) Nasopharyngeal carriage of *Streptococcus pneumoniae* and pneumococcal urine antigen test in healthy elderly subjects. *Scand J Infect Dis* **44**, 433–8.

[25] Perrocheau, A, Doyle, A, Bernillon, P, & Varon, E. e. a. (2006) Estimation du nombre total de méningites à pneumocoque de l'enfant par la méthode capture-recapture à 3 sources, France, 2001-2002. *Bull Epidemiol Hebd* pp. 2–3.

[26] Kono, M, Zafar, M. A, Zuniga, M, Roche, A. M, Hamaguchi, S, & Weiser, J. N. (2016) Single cell bottlenecks in the pathogenesis of *Streptococcus pneumoniae*. *PLoS Pathog* **12**, e1005887.

[27] Weinberger, D. M, Grant, L. R, Steiner, C. A, Weatherholtz, R, Santosham, M, Viboud, C, & O'Brien, K. L. (2014) Seasonal drivers of pneumococcal disease incidence: impact of bacterial carriage and viral activity. *Clin Infect Dis* **58**, 188–94.

[28] Eames, K. T. D, Tilston, N. L, Brooks-Pollock, E, & Edmunds, W. J. (2012) Measured dynamic social contact patterns explain the spread of h1n1v influenza. *PLoS Comput Biol* **8**, e1002425.

[29] Le Polain de Waroux, O, Flasche, S, Prieto-Merino, D, & Edmunds, W. J. (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**, e86136.

[30] Krone, C. L, van de Groep, K, Trzciński, K, Sanders, E. A. M, & Bogaert, D. (2014) Immunosenescence and pneumococcal disease: an imbalance in host-pathogen interactions. *Lancet Respir Med* **2**, 141–53.

[31] King, A. A, Nguyen, D, & Ionides, E. L. (2016) Statistical inference for partially observed Markov processes via the R package pomp. *Journal of Statistical Software* **69**, 1–43.

[32] Zucchini, W & MacDonald, I. L. (2009) *Hidden Markov models for time series: an introduction using R*, Monographs on statistics and applied probability. (CRC Press, Boca Raton) Vol. 110.

[33] Johnson, S. G. (?) The nlopt nonlinear-optimization package. *?* **?**, ?

[34] R Core Team. (2016) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).

[35] Microsoft Corporation. (2017) *checkpoint: Install Packages from Snapshots on the Checkpoint Server for Reproducibility.* R package version 0.4.0.

[36] Rohani, P, Zhong, X, & King, A. A. (2010) Contact network structure explains the changing epidemiology of pertussis. *Science* **330**, 982–5.

[37] Domenech de Cellès, M, Arduin, H, Varon, E, Souty, C, Boëlle, P.-Y, Lévy-Bruhl, D, van der Werf, S, Soulary, J.-C, Guillemot, D, Watier, L, & Opatowski, L. (2018) Characterizing and comparing the seasonality of influenza-like illnesses and invasive pneumococcal diseases using seasonal waveforms. *Am J Epidemiol* **187**, 1029–1039.

[38] Numminen, E, Chewapreecha, C, Turner, C, Goldblatt, D, Nosten, F, Bentley, S. D, Turner, P, & Corander, J. (2015) Climate induces seasonality in pneumococcal transmission. *Sci Rep* **5**, 11344.

[39] Bojang, A, Jafali, J, Egere, U. E, Hill, P. C, Antonio, M, Jeffries, D, Greenwood, B. M, & Roca, A. (2015) Seasonality of pneumococcal nasopharyngeal carriage in rural gambia determined within the context of a cluster randomized pneumococcal vaccine trial. *PLoS One* **10**, e0129649.

[40] Usuf, E, Bottomley, C, Adegbola, R. A, & Hall, A. (2014) Pneumococcal carriage in sub-Saharan Africa–a systematic review. *PLoS One* **9**, e85001.

[41] Adegbola, R. A, DeAntonio, R, Hill, P. C, Roca, A, Usuf, E, Hoet, B, & Greenwood, B. M. (2014) Carriage of *Streptococcus pneumoniae* and other respiratory bacterial pathogens in low and lower-middle income countries: a systematic review and meta-analysis. *PLoS One* **9**, e103293.

[42] Weinberger, D. M, Klugman, K. P, Steiner, C. A, Simonsen, L, & Viboud, C. (2015) Association between respiratory syncytial virus activity and pneumococcal disease in infants: a time series analysis of US hospitalization data. *PLoS Med* **12**, e1001776.

[43] Paireau, J, Pelat, C, Caserio-Schönemann, C, Pontais, I, Le Strat, Y, Lévy-Bruhl, D, & Cauchemez, S. (2018) Mapping influenza activity in emergency departments in France using Bayesian model-based geostatistics. *Influenza Other Respir Viruses* **12**, 772–779.