Supplementary Information Appendix - Increased transmissibility explains the third wave of infection by the 2009 H1N1 pandemic virus in England

Ilaria Dorigatti[1], Simon Cauchemez and Neil M. Ferguson

July 21, 2013

MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, Norfolk Place, London W2 1PG, United Kingdom
1 Author for correspondence: i.dorigatti@imperial.ac.uk
phone: +44 (0)20 7594 3229

Contents

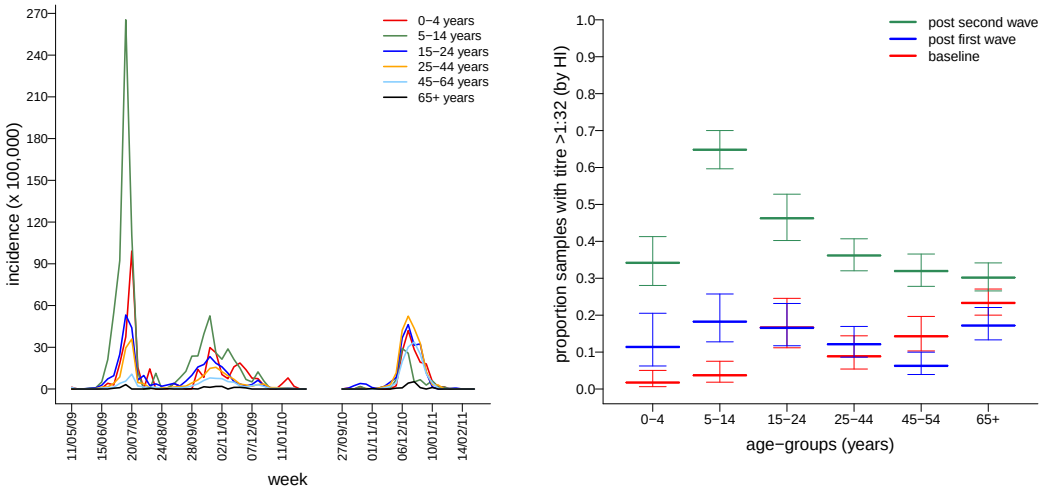
1 Materials and Methods

1.1 Data

1.1.1 Syndromic, virological and serological data

The syndromic data used in this work have been collected by QSurveillance, a syndromic surveillance scheme that monitors general practitioners (GP) consultations in the UK since 2004 (`http://www.qsurveillance.org`). Syndromic data consist of the weekly numbers of GP consultations due to Influenza Like Illness (ILI) symptoms together with the weekly size of the monitored patient population.
During the 2009 pandemic, the Royal College of General Practitioners (RCGP) (`http://www.rcgp.org.uk`) and HPA Regional Microbiology Network (RMN) coordinated a swabbing program where patients diagnosed with ILI were tested by Polymerase Chain Reaction (PCR) for H1N1pdm09 influenza virus. We aggregated the RCGP and HPA-RMN data by week of sample collection, and age group, and analysed the numbers of samples testing positive among those tested.
During 2009-2010, the HPA conducted a serological study to determine the levels of immunity to H1N1pdm09 virus existing in the English population prior the pandemic and to investigate the changes in seroprevalence as the epidemic unfolded. Cross-sectional surveys of seroprevalence were undertaken at a pre-pandemic baseline (using samples collected in 2008), after the first wave (using samples collected in August 2009) and after the second wave (using samples collected from January to April 2010) [1]. Serum samples were collected through the HPA-RMN at the baseline and additional samples were provided by chemical pathology laboratories during the pandemic [1]. The samples were tested by haemagglutination inhibition (HI) and the results, in terms of number of samples showing HI titres over $1:32$ among those tested by age group at the pre-pandemic baseline, after the first wave and after the second wave are presented in [1]. Sample numbers showed geographical biases, so we generated a unique synthetic dataset from the reported "London" and "outside London" aggregations of the data by taking the population weighted average of the two datasets (see Section 1.1.3 below).
In Figure S1 we plot the H1N1-attributable ILI incidence curves obtained by multiplying weekly ILI incidence in an age group by the proportion of swabs testing positive that week in that age class. Figure S1 also shows the seroprevalence by age group (0-4, 5-14, 15-24, 25-44, 65+ years).

Figure S1: Plot of surveillance data. Left panel: Incidence (per 100,000) of the number of reported H1N1-attributable ILI cases by age group (0-4, 5-14, 15-24, 25-44, 45-64, 65+ years), obtained by multiplying the weekly ILI datum by the proportion of positive samples on the corresponding age group and week (data sources: QSurveilance, RCGP, HPA-RMN). Right panel: Plot of the proportion of samples with HI tires $> 1:32$ by age group (0-4, 5-14, 15-24, 25-44, 45-64, 65+ years) at the pre-pandemic baseline, after the first wave and after the second wave according to the synthetic serological dataset for England.



1.1.2 National Pandemic Flu Service

In order to relieve the high demand on GPs during the first wave, the UK Department of Health activated the National Pandemic Flu Service (NPFS), a web and telephone service to support the assessment of patients' symptoms and their need for antivirals. NPFS was restricted to individuals older than 1 year and with no serious underlying illness or pregnancy; infants and people in risk groups continued to be referred to to their GP. The service was launched in England on 23 July 2009, at the peak of the first wave. Given the low levels of influenza activity in the population at the end of the second wave, NPFS was withdrawn on 11 February 2010. In our analysis we do not use data from NPFS.

1.1.3 Synthetic serological dataset

We describe here the method used to generate the synthetic serological dataset representative for all England from the data published in [1]. In [1], samples were collected from all English regions except East Midlands and the results of the analysis were given for "London" and "outside London".

Let pop_L^i denote the London population of age group $i = 1, \ldots, 6$ and pop_N^i denote the population of England except London and East Midlands of age group $i = 1, \ldots, 6$. Let X_{Lt}^i and Y_{Lt}^i denote the number of serum samples collected in age group i at time t and the number of samples showing HI titre $> 1:32$ in age group i at time t in London respectively. Similarly, let X_{Nt}^i and Y_{Nt}^i denote the number of samples collected in age group i at time t and the number of samples showing HI titre $> 1:32$ in age group i at time t outside London respectively. In our synthetic dataset the denominator, i.e. the number of samples tested X_t^i is given by the sum of the samples tested in London and outside London $X_t^i = X_{Lt}^i + X_{Nt}^i$.

We generated a synthetic numerator Y_t^i by taking the population weighted average of the fractions showing HI titre $> 1:32$

$$Y_t^i = \left(\frac{{Y_L}_t^i}{{X_L}_t^i} \frac{\text{pop}_L^i}{\text{pop}_L^i + \text{pop}_N^i} + \frac{{Y_N}_t^i}{{X_N}_t^i} \frac{\text{pop}_N^i}{\text{pop}_L^i + \text{pop}_N^i} \right) X_t^i \tag{1}$$

We use X_t^i and Y_t^i as serological data for the whole England.

1.2 Data uncertainty

We calculate exact 95% confidence intervals (CI) around the observed data. The 95% CI around the observed proportions of samples testing seropositive (see right panel of Figure S1) has been computed by assuming that samples with titres $> 1:32$ are binomial distributed. The 95% CI around the observed H1N1 attributable ILI data (see for instance Figure 2 in the main text and Fig. S4-S9) has been computed assuming that ILI data are Poisson distributed and that the number of positive swabs among those tested is binomial distributed.

1.3 Vaccination

When the pandemic was declared, on 11 June 2009, the UK government ordered enough vaccine to cover 100% of the population with two doses [2]. Despite this, vaccine supplies were initially limited and the pandemic vaccination program, which started in October 2009, targeted at-risk groups [3]. Data show that vaccination coverage remained below 5% in all age groups up to week 48 of 2009 (i.e. after the peak of the second wave) [1]. Overall, at the end of 2009, the pandemic vaccine uptake was $< 5\%$ in under 45 year olds, about 10% in $45-64$ year olds and about 20% in 65+ year olds [1]. From January 2010 (i.e. at the end of the second wave) children under 5 were targeted by an extensive vaccination program which reached up to 30% coverage [5], while the percentage of population vaccinated with the pandemic strain vaccine did not significantly increase in the other age groups [1]. In April 2010, given the low level of activity of H1N1pdm09, the Department of Health reached an agreement with the vaccine manufacturers to only take deliveries of 35 million doses of pandemic vaccine [2].
In 2010-2011, seasonal vaccination of over 65 started in October 2010 (i.e. during the third wave) and reached 72.8% coverage [4].

1.4 The model

1.4.1 Transmission model

The age-structured deterministic transmission model described in the main text is defined by the following differential equations

$$\begin{aligned}
\dot{S}_i &= \delta_1 P S_i - \lambda_i(t) S_i \\
\dot{PS}_i &= -\delta_1 P S_i - \sigma \lambda_i(t) P S_i \\
\dot{I}_i &= \lambda_i(t)(\sigma P S_i + S_i) - \gamma I_i \\
\dot{R^-}_i &= \gamma I_i - \omega R_i^- \\
\dot{R^+}_i &= \omega R_i^- .
\end{aligned} \tag{2}$$

where $i = 1, \ldots, 6$ indexes the six age groups 0-4, 5-14, 15-24, 25-44, 45-64, 65+ years respectively. The flow diagram of model (2) is given in Figure S2

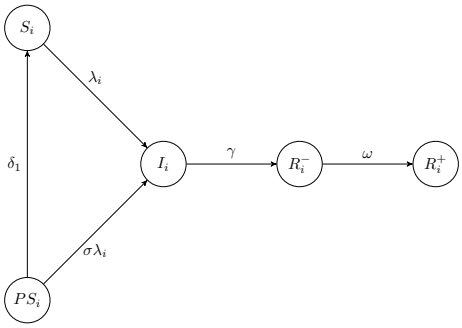


Figure S2: Flow diagram describing the transmission model. We assume that at the start of the pandemic individuals are either fully susceptible (S) to infection by H1N1pdm09 or have some degree of prior immunity against infection by H1N1pdm09, in which case they are classified as partially susceptible (PS). After infection, individuals move to the infectious state (I) and eventually recover. We model the biological delay to seroconversion by splitting the class of recovered individuals into two sub-classes: R^- denotes the recovered and fully immune testing seronegative, R^+ denotes the recovered and fully immune testing seropositive. Parameter σ represents the susceptibility of class PS, λ represents the force of infection, γ represents the recovery rate, ω represents the rate of seroconversion and δ_1 represents the rate of decay of prior immunity. Subscript i denotes the age-class ($i = 1, \ldots, 6$).

The force of infection λ_i is given by

$$\lambda_i(t) = p(t) \sum_{j=1}^{6} c_{ij} \frac{I_j(t)}{N_j} \quad i = 1, \ldots, 6 \tag{3}$$

where $p(t)$ is the probability of getting infected upon a contact with an infectious individual, c_{ij} indicates the mean number of contacts between an individual of age class i with individuals of age class j and N_i represents the constant in time size of age group i. We assume that $p(t)$ is piecewise constant in each wave

$$p(t) = \begin{cases} p_1 & \text{if } t \in \text{weeks [16/2009, 36/2010) (wave 1)} \\ p_2 & \text{if } t \in \text{weeks [36/2009, 38/2010) (wave 2)} \\ p_3 & \text{if } t \in \text{weeks [38/2010, 8/2011) (wave 3)} \end{cases} \tag{4}$$

We assume that at the beginning of the pandemic a fraction of the population in each age group has some prior immunity to infection by H1N1pdm09, due to previous exposure to other influenza viruses or vaccination, while the rest of the population is completely naive. Individuals with prior immunity are classified as partially susceptible (PS). We let $ps_b^i = PS_i(t)/N_i$ denote the fraction of partially susceptible population in age group i ($i = 1, \ldots, 6$) at the baseline (i.e. at week 26/2008, where we start the simulation - see paragraph 1.5.2). The effective reproduction number R_e is the expected number of cases generated by a single infective when a fraction of the population is (partially) protected against infection and is given by

$$R_e = \frac{p(t)s(K)}{\gamma} \tag{5}$$

where $s(K)$ denotes the spectral radius of matrix K [6]. Matrix K is given by

$$k_{ij}(t) = c_{ji}\Big(\sigma\frac{PS_i(t)}{N_i} + \frac{S_i(t)}{N_i}\Big) \qquad i,j = 1,\ldots,6. \tag{6}$$

The transmission model outputs C_t^i, the number of A/H1N1 infections in week t and age group i ($i = 1,\ldots,6$) in the English population. By scaling C_t^i down to the size of the monitored population (*i.e.* that covered by QSurveillance), we obtain Z_t^i, the expected number of A/H1N1 infections generated within class i during week t in the monitored population. This model was coded in C and numerically solved using the Runge-Kutta algorithm with a fixed step size of 0.05 days [7].

1.4.2 Statistical model

Likelihood

Let ILI_t^i denote the number of ILI cases in the monitored population of age group i and week t and ρ_t^i represent the probability that an individual in age group i infected with H1N1pdm09 reports ILI symptoms in week t. Let P_t^i represent the number of positive swabs among those tested in age group i at week t and T_t^i denote the number of swabs tested in age group i at week t. Finally, let X_t^i represent the number of blood samples tested in age group i at week t and Y_t^i represent the number of seropositive samples among those tested in age class i at time t.
Using θ to denote the parameter vector, the posterior likelihood is defined by

$$P(\{ILI_t^i\}_{i,t}, \{P_t^i\}_{i,t}, \{Y_t^i\}_{i,t}, \theta | T_t^i, X_t^i) =$$

$$\prod_t \prod_i P(ILI_t^i, P_t^i | T_t^i, Z_t^i(\theta), \rho_t^i) P(Y_t^i | X_t^i, \theta) P(\theta) \tag{7}$$

where $P(ILI_t^i, P_t^i | T_t^i, Z_t^i(\theta), \rho_t^i)$ denotes the probability of syndromic and virological data, $P(Y_t^i | X_t^i, \theta)$ denotes the probability of serological data and $P(\theta)$ denotes the prior distribution. In the following paragraphs we derive the explicit formulation for the probability of syndromic and virological data (see eq. (15)) and for the probability of serological data (see eq. (16)).

Probability of syndromic and virological data

We adopt the modelling framework developed in [8] and here we describe the main assumptions leading to the formulation of a probabilistic model for the ILI and virological data. A more complete discussion is given in [8].
Let π_t^i represent the probability of testing positive in age group i ($i = 1,\ldots,6$) at week t. Following [8], we assume that the distribution of the number of positive swabs P_t^i is binomial with parameters T_t^i and π_t^i

$$P(P_t^i | T_t^i, \pi_t^i) = \binom{T_t^i}{P_t^i} {\pi_t^i}^{P_t^i} (1-\pi_t^i)^{T_t^i - P_t^i} \tag{8}$$

and that the prior distribution of π_t^i is Beta with parameters α and β

$$P(\pi_t^i) = \frac{{\pi_t^i}^{\alpha-1}(1-\pi_t^i)^{\beta-1}}{B(\alpha,\beta)} \tag{9}$$

Let F_t^i represent the number of individuals with H1N1pdm09 infection in age group i reporting ILI symptoms in week t. We assume that the distribution of F_t^i is binomial with parameters Z_t^i and ρ_t^i

$$P(F_t^i|Z_t^i,\rho_t^i) = \binom{Z_t^i}{F_t^i}{\rho_t^i}^{F_t^i}(1-\rho_t^i)^{Z_t^i-F_t^i} \tag{10}$$

If $F_t^i > 0$, we assume that the number of ILI cases in the monitored population of age group i and week t has a negative binomial distribution with parameters F_t^i and π_t^i

$$P(ILI_t^i|F_t^i,\pi_t^i) = \binom{ILI_t^i-1}{F_t^i-1}{\pi_t^i}^{F_t^i}(1-\pi_t^i)^{ILI_t^i-F_t^i}. \tag{11}$$

If $F_t^i = 0$, then the whole ILI_t^i is uninfected with H1N1pdm09, that is

$$P(ILI_t^i|F_t^i,\pi_t^i) = (1-\pi_t^i)^{ILI_t^i}. \tag{12}$$

Using conditional probability we compute the probability of syndromic and virological data given the model

$$P(ILI_t^i, P_t^i|T_t^i, Z_t^i, \rho_t^i) =$$

$$\sum_{j=1}^{\min(ILI_t^i, Z_t^i)} P(ILI_t^i|F_t^i=j,\pi_t^i)P(F_t^i=j|Z_t^i,\rho_t^i)P(P_t^i|T_t^i,\pi_t^i). \tag{13}$$

By integration over π_t^i

$$P(ILI_t^i, P_t^i|T_t^i, Z_t^i, \rho_t^i) =$$

$$\sum_{j=1}^{\min(ILI_t^i, Z_t^i)} \int_0^1 P(ILI_t^i|F_t^i=j,\pi_t^i)P(F_t^i=j|Z_t^i,\rho_t^i)P(P_t^i|T_t^i,\pi_t^i)P(\pi_t^i)d\pi_t^i \tag{14}$$

we obtain the following explicit formula for the probability of syndromic and virological data given the model (see [8] for details)

$$P(ILI_t^i, P_t^i | T_t^i, Z_t^i, \rho_t^i) =$$

$$\frac{\binom{T_t^i}{P_t^i}}{B(\alpha,\beta)}\Bigg((1-\rho_t^i)^{Z_t^i} B(P_t^i+\alpha, ILI_t^i+T_t^i-P_t^i+\beta)+$$

$$+\sum_{F_t^i=1}^{\min(ILI_t^i,Z_t^i)} \binom{ILI_t^i-1}{F_t^i-1}\binom{Z_t^i}{F_t^i}(\rho_t^i)^{F_t^i}(1-\rho_t^i)^{Z_t^i-F_t^i} B(F_t^i+P_t^i+\alpha, ILI_t^i-F_t^i+T_t^i-P_t^i+\beta)\Bigg). \tag{15}$$

Probability of serological data

By definition, individuals in the model compartment R^+ have an antibody titre $> 1:32$ (by HI). We assume that partially susceptible individuals have a certain probability ξ of having HI titre $> 1:32$ and that individuals in the other infectious states (S, I and R^-) are all seronegative (i.e. have HI titres $< 1:32$). We assume that the distribution of Y_t^i is binomial with parameters X_t^i and w_t^i

$$P(Y_t^i | X_t^i, w_t^i) = \binom{X_t^i}{Y_t^i} \left(w_t^i\right)^{Y_t^i} \left(1-w_t^i\right)^{X_t^i-Y_t^i} \tag{16}$$

where the probability of testing seropositive is given by

$$w_t^i = \xi \frac{PS_i(t)}{N_i} + \frac{R_i^+(t)}{N_i} \tag{17}$$

According to the timing of collection of the serum samples described in [1], we assume that seroprevalence is measured in the population at week 26/2008 for pre-pandemic baseline, at week 36/2009 for post-first wave serology and at week 14/2010 for post-second wave serology. Given that the H1N1pdm09 virus had not emerged at baseline, at week $t = 26/2008$ the probability of testing seropositive reduces to $w_t^i = \xi ps_b^i$ where for simplicity of notation we denoted the pre-pandemic baseline fraction of partially susceptible individuals in age group i with $ps_b^i = PS_i(t)/N_i$.

1.5 Model parameterization

1.5.1 Reporting rates

We allow for people in different age groups having different propensities to consult their GP for ILI and that the reporting rate may have varied in time.
Let $\mathrm{AR}^i_{\mathrm{sero}}$ denote the serological attack rate in age group i in waves 1 and 2 and $\mathrm{AR}^i_{\mathrm{ILI\&H1N1}}$ denote the cumulative incidence of H1N1-attributable ILI cases (obtained by multiplying the weekly ILI incidence by the proportion of positive samples in that week for each age group) in age group i in waves 1 and 2. We assume the ratio of the syndromic to serological attack

rate

$$\phi^i = \frac{\mathrm{AR}^i_{\mathrm{ILI\&H1N1}}}{\mathrm{AR}^i_{\mathrm{sero}}} \tag{18}$$

quantifies the propensity of age group i to consult their GP for ILI symptoms. We further assume that temporal changes in reporting affected all age groups equally, so that the reporting rate of age group i at time t can be expressed as a time dependent baseline reporting r_t multiplied by an age-dependent component ϕ^i

$$\rho^i_t = \phi^i r_t \quad i = 1, \ldots, 6 \tag{19}$$

We take ρ^4_t, the reporting rate of age group 4 (25-44 years) as the reference reporting rate to be estimated from the data. From (19) it follows that the reporting rate of age group i can be expressed as

$$\rho^i_t = \frac{\phi^i}{\phi^4}\rho^4_t = \psi^i \rho^4_t \quad i = 1, \ldots, 6 \tag{20}$$

From the data we compute the following values for ψ^i:

$$\psi^1 = 1.57, \quad \psi^2 = 1.97, \quad \psi^3 = 1.47, \quad \psi^5 = 0.71, \quad \psi^6 = 0.30.$$

We finally assume that the reporting rate is piecewise constant in the time windows from the beginning of the epidemic to the week of introduction of NPFS (week 30/2009), from the introduction of NPFS to the end of the second wave and in the third wave:

$$\rho^4_t = \begin{cases} \rho^4_1 & \text{if } t \in \text{weeks [16/2009, 31/2009)} \\ \rho^4_2 & \text{if } t \in \text{weeks [31/2009, 05/2010)} \\ \rho^4_3 & \text{if } t \in \text{weeks [38/2010, 08/2011)} \end{cases}$$

Estimation of the age- and time-dependent reporting rates is thereby reduced to the estimation of ρ^4_1, ρ^4_2, ρ^4_3.

1.5.2 School holidays and infection seeding

In England, during the time window commencing on week 20/2009 and ending on week 5/2010, schools closed on week 22/2009 for the mid-term spring break, from week 30/2009 to week 35/2009 for the summer vacation, on week 44/2009 for the mid-term fall break and on weeks 52/2009 and 1/2010 for the Christmas holidays. During the time window from week 40/2010 to week 7/2011, schools closed from 20/12/2010 (start of week 52/2010) to 04/01/2011 (second day of week 2/2011) for Christmas.
We model mid-term (spring and fall) breaks and Christmas holidays by rescaling the contacts of school-age children with school-age children ($5 - 14$ years) by a factor ϵ_1 and by rescaling the contacts of all other age groups by a factor ϵ_2. If c_{ij} indicates the mean number of contacts between an individual of age class i with individuals of age class j ($i, j = 1, \ldots, 6$), we assume

that during the mid-term breaks and Christmas, the mean number of contacts is given by

$$\begin{pmatrix} \epsilon_2 c_{11} & \epsilon_2 c_{12} & \epsilon_2 c_{13} & \cdots & \epsilon_2 c_{16} \\ \epsilon_2 c_{21} & \epsilon_1 c_{22} & \epsilon_2 c_{13} & \cdots & \epsilon_2 c_{26} \\ \epsilon_2 c_{31} & \epsilon_2 c_{32} & \epsilon_2 c_{33} & \cdots & \epsilon_2 c_{36} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon_2 c_{61} & \epsilon_2 c_{62} & \epsilon_2 c_{63} & \cdots & \epsilon_2 c_{66} \end{pmatrix} \tag{21}$$

Similarly, to reproduce summer school holidays, we scale the contact rate of school-aged children with school-aged children ($5-14$ years) by a factor ϵ_3 and rescale all other contact rates by a different reduction factor ϵ_4.
We start the simulations at week 26/2008, the time we assume the baseline serology was measured, seed an initial number of H1N1 cases, I_1, on week 17/2009 (late April 2009) and fit the model to surveillance data in the time window between week 20/2009 (early May 2009) and week 5/2010 (end of January 2010). To trigger the third wave, we re-seed a number I_3 of H1N1 cases, on week 37/2010 (early September 2010) and fit the model to surveillance data in the time window between week 40/2010 (end of September 2010) and week 7/2011 (beginning of February 2011). No influenza activity was recorded between these two time windows (i.e. from week 6/2010 to week 39/2010). The initial numbers of cases I_1, I_3 are distributed among the age classes in proportions 5%, 40%, 30%, 10%, 10% and 5%, respectively, though results were not sensitive to this choice.

1.5.3 Summary of parameters used

A summary of the parameters used in the model is given in Table S1.

Table S1: Summary of model parameter values (excluding contact rates)

Notation	Parameter	Value
R_e	effective reproduction number	estimated
p_1	probability of transmission in wave 1	estimated
p_2	probability of transmission in wave 2	fixed or estimated
p_3	probability of transmission in wave 3	fixed or estimated
δ_1	rate of decay of prior immunity ($PS \to S$)	fixed or estimated
γ	infectious rate	0.345/day
ω	seroconversion rate	0.071/day
ρ_1^4	reporting rate of age group 4 in weeks $[16/2009, 31/2009)$	estimated
ρ_2^4	reporting rate of age group 4 in weeks $[31/2009, 05/2010)$	estimated
ρ_3^4	reporting rate of age group 4 in weeks $[38/2010, 08/2011)$	estimated
ψ^1	propensity of age group 1 compared to age group 4 to consult GP due to ILI	1.57
ψ^2	propensity of age group 2 compared to age group 4 to consult GP due to ILI	1.97
ψ^3	propensity of age group 3 compared to age group 4 to consult GP due to ILI	1.47
ψ^5	propensity of age group 5 compared to age group 4 to consult GP due to ILI	0.71
ψ^6	propensity of age group 6 compared to age group 4 to consult GP due to ILI	0.30
ξ	probability of showing HI titre $> 1:32$ if in prior immune compartment (PS)	estimated
I_1	initial number of infections seeding wave 1 on week 17/2009	estimated
I_3	initial number of infections seeding wave 3 on week 37/2010	estimated
σ	susceptibility of partially susceptible (PS)	estimated
ps_b^i	baseline fraction of individuals with prior immunity in age group $i (i = 1, \ldots, 6)$	estimated
ϵ_1	reduction contacts/day within age group 2 in mid-term and Christmas holidays	estimated
ϵ_2	as ϵ_1 but in all age groups except within age group 2	estimated
ϵ_3	reduction contacts/day within age group 2 in summer holidays	estimated
ϵ_4	as ϵ_3 but in all age groups except within age group 2	estimated

1.6 Assessing model fit

ILI and virology

For each model variant, we draw 500 parameter sets from the posterior distribution. For each parameter set, we numerically solve the deterministic transmission model and rescale the solution to the monitored patient population size, thus obtaining Z_t^i, the expected number of patients infected by H1N1pdm09 in age class i at time t. By multiplying Z_t^i by the estimated reporting rate ρ_t^i, we obtain the expected number of H1N1pdm09 infections reporting ILI in age class i at time t.

Serology

The transmission model also provides estimates of the fraction of individuals in states PS and R^+ in age group i at time t. We can hence compute the expected proportion of the population testing seropositive at weeks $t = 26/2008$ (representing the baseline serology), $t = 36/2009$ (representing the post-first wave serology) and $t = 14/2010$ (representing the post-second wave serology) by adding the fraction of population with homologous immunity (R^+) to the fraction of population with partial prior protection (PS) multiplied by the estimated probability that someone with prior immunity will test seropositive ξ, according to equation (17).

2 Results

2.1 Fit of models 1-6 to the 2009-2011 ILI consultation and virological data and to the 2009-2010 serological data

We summarize in Table S2 the DIC score, the posterior mean and 95% credible interval (CrI) of the log-likelihood and parameters for model variants 1-6. For the sake of clarity, instead of reporting the estimated probability of transmission in waves 2 and 3 (i.e. p_2 and p_3), we present the estimated relative change in the probability of transmission in waves 2 and 3 compared to wave 1 (i.e. p_2/p_1 and p_3/p_1). Similarly, instead of presenting the estimated reporting rate in each wave, we show the estimated relative change in the reporting rate in waves 2 and 3 compared to wave 1 (i.e. ρ_2^4/ρ_1^4 and ρ_3^4/ρ_1^4). We also present the half-life of prior immunity rather than the decay rate:

$$t_{1/2}(\delta_1) = -\frac{1}{\delta_1}\frac{1}{365}\ln\left(\frac{1}{2}\right).$$

In Figure S3 we show the DIC of models 1-6. In Figures S4-S9 we present the fit to the H1N1-attributable ILI incidence curves and to the pre-pandemic baseline, post-first wave and post second wave serology obtained with models 1-6.
Figure S10 shows the predicted seroprevalence (by age group) at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09), after the second wave (week 14/10), before the third wave (week 44/10) and after the third wave (week 18/11) obtained with model 6 calibrated on the 2009-2010 serology.
Figure S11 shows the fit to the H1N1-attributable ILI and seroprevalence data obtained with model 6 in its variant with vaccination (see Materials and methods in the main text for details on modelling of vaccination). Figure S12 shows the estimated effective reproduction number in time, incidence of H1N1 infections in the population and population attack rates by age group and wave obtained with model variant 6 when vaccination is included.

Table S2: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters for model variants 1-6

Model	1	2	3	4	5	6
DIC	15830.5	15814.7	8337.5	8282.5	8258.4	8172.1
Log-like.	-7913 (-7920, -7908)	-7907 (-7914, -7902)	-4169 (-4177, -4164)	-4150 (-4158, -4144)	-4135 (-4144, -4129)	-4118 (-4126, -4112)
R_e	1.44 (1.41, 1.46)	1.41 (1.38, 1.45)	1.45 (1.43, 1.47)	1.461 (1.43, 1.49)	1.43 (1.41, 1.45)	1.41 (1.39, 1.43)
p_1	0.052 (0.049, 0.056)	0.051 (0.048, 0.055)	0.051 (0.048, 0.054)	0.057 (0.052, 0.062)	0.052 (0.050, 0.054)	0.051 (0.049, 0.052)
p_2/p_1	- -	- -	- -	1.15 (1.10, 1.20)	- -	1.08 (1.05, 1.11)
p_3/p_1	- -	- -	1.92 (1.89, 1.95)	2.21 (2.11, 2.32)	1.67 (1.62, 1.72)	1.70 (1.64, 1.76)
$t_{1/2}(\delta_1)$ (years)	- -	1.79 (0.96, 9.25)	- -	- -	1.55 (1.40, 1.68)	1.03 (0.95, 1.16)
ρ_1^4	0.0142 (0.0123, 0.0162)	0.0125 (0.0107, 0.0147)	0.0137 (0.0121, 0.0156)	0.0084 (0.0070, 0.0100)	0.0122 (0.0108, 0.0135)	0.0081 (0.0069, 0.0094)
ρ_2^4/ρ_1^4	0.31 (0.26, 0.36)	0.35 (0.30, 0.40)	0.31 (0.27, 0.35)	0.57 (0.46, 0.69)	0.34 (0.30, 0.39)	0.53 (0.45, 0.61)
ρ_3^4/ρ_1^4	0.46 (0.02, 0.97)	0.40 (0.021, 0.93)	0.23 (0.19, 0.26)	0.39 (0.32, 0.48)	0.28 (0.24, 0.32)	0.47 (0.40, 0.56)
ξ	0.21 (0.18, 0.25)	0.24 (0.20, 0.29)	0.18 (0.15, 0.21)	0.15 (0.13, 0.18)	0.20 (0.17, 0.23)	0.20 (0.17, 0.23)
I_1	2.2 (1.3, 3.8)	3.1 (1.6, 5.3)	1.6 (1.0, 2.7)	2.8 (1.4, 5.1)	1.9 (1.1, 3.2)	5.1 (2.6, 7.8)
I_3	70.4 (4.5, 208.6)	42.3 (4.4, 89.0)	1.7 (1.0, 3.1)	1.5 (1.0, 2.5)	1.8 (1.0, 3.6)	1.7 (1.0, 3.2)
σ	0.41 (0.30, 0.52)	0.26 (0.09, 0.41)	0.57 (0.49, 0.63)	0.48 (0.42, 0.54)	0.39 (0.32, 0.46)	0.24 (0.19, 0.30)
ps_b^1	0.08 (0.02, 0.17)	0.08 (0.02, 0.17)	0.20 (0.07, 0.33)	0.42 (0.19, 0.60)	0.13 (0.03, 0.33)	0.09 (0.03, 0.20)
ps_b^2	0.16 (0.07, 0.25)	0.14 (0.06, 0.23)	0.13 (0.06, 0.22)	0.15 (0.06, 0.27)	0.12 (0.05, 0.21)	0.07 (0.03, 0.11)
ps_b^3	0.56 (0.43, 0.70)	0.63 (0.45, 0.83)	0.79 (0.62, 0.96)	0.84 (0.73, 0.94)	0.84 (0.70, 0.98)	0.78 (0.66, 0.91)
ps_b^4	0.56 (0.42, 0.72)	0.66 (0.45, 0.90)	0.73 (0.54, 0.94)	0.89 (0.75, 0.99)	0.91 (0.76, 0.99)	0.94 (0.85, 1.00)
ps_b^5	0.49 (0.36, 0.63)	0.56 (0.38, 0.77)	0.59 (0.43, 0.77)	0.79 (0.63, 0.93)	0.74 (0.57, 0.92)	0.84 (0.68, 0.98)
ps_b^6	0.95 (0.83, 1.00)	0.94 (0.81, 1.00)	0.98 (0.93, 1.00)	0.98 (0.95, 1.00)	0.98 (0.92, 1.00)	0.97 (0.91, 1.00)
ϵ_1	0.70 (0.36, 0.95)	0.64 (0.25, 0.94)	0.04 (0.00, 0.12)	0.06 (0.00, 0.21)	0.03 (0.00, 0.11)	0.04 (0.00, 0.12)
ϵ_2	0.99 (0.96, 1.00)	0.99 (0.96, 1.00)	0.99 (0.98, 1.00)	0.99 (0.98, 1.00)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)
ϵ_3	0.36 (0.17, 0.51)	0.32 (0.13, 0.49)	0.48 (0.33, 0.60)	0.55 (0.42, 0.66)	0.41 (0.25, 0.54)	0.43 (0.25, 0.57)
ϵ_4	0.65 (0.61, 0.69)	0.66 (0.62, 0.69)	0.61 (0.57, 0.65)	0.66 (0.60, 0.70)	0.63 (0.59, 0.67)	0.65 (0.61, 0.69)

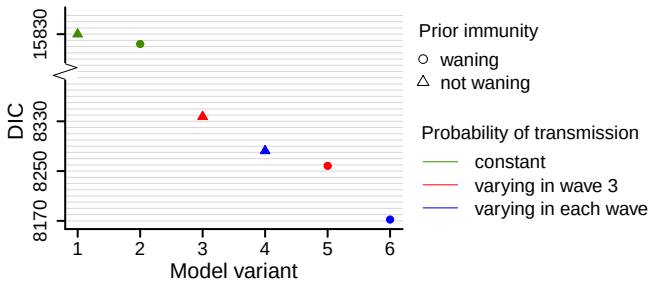


Figure S3: Deviance Information Criterion (DIC) score of models 1-6.

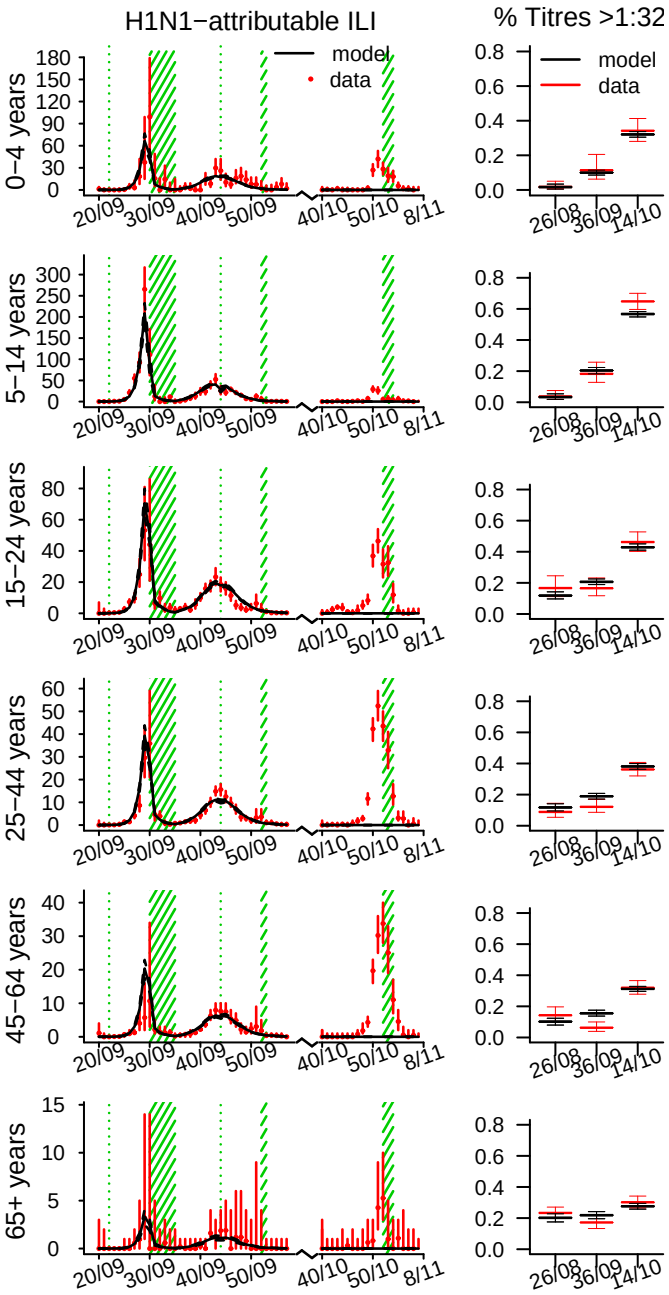


Figure S4: Age specific illness rates and seroprevalence estimated using model variant 1. Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% CI around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1:32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09) and after the second wave (week 14/10).

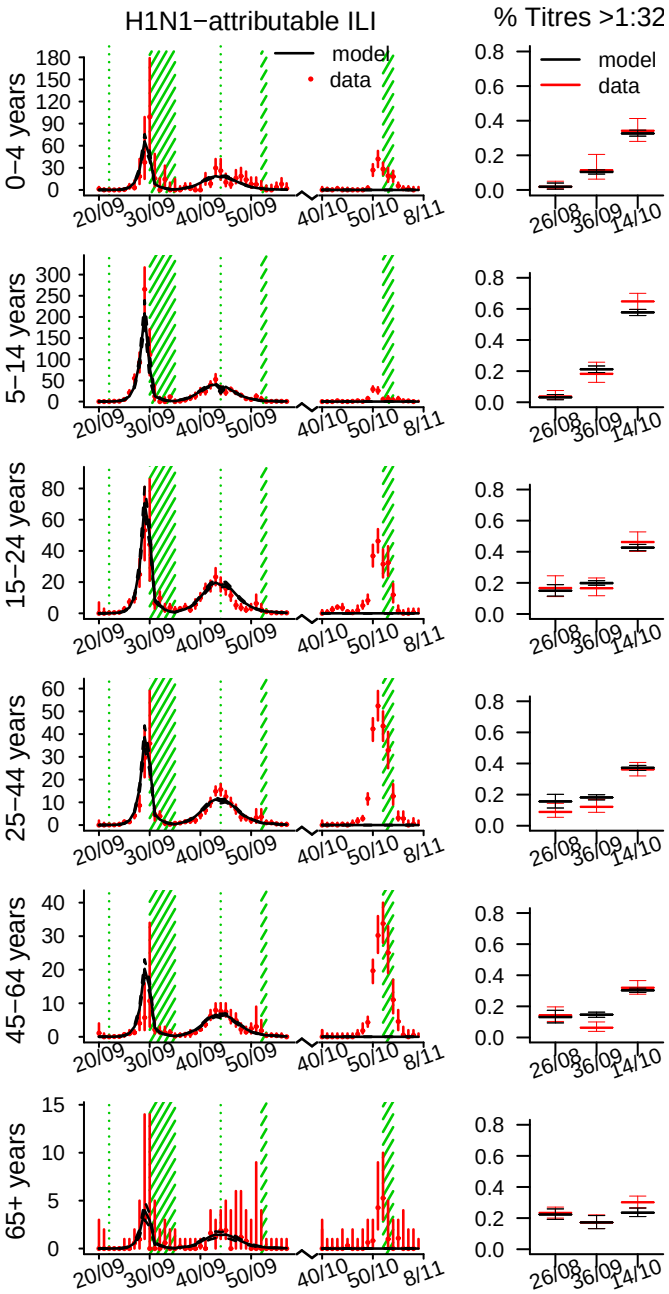


Figure S5: As Figure S4, but obtained using model variant 2.

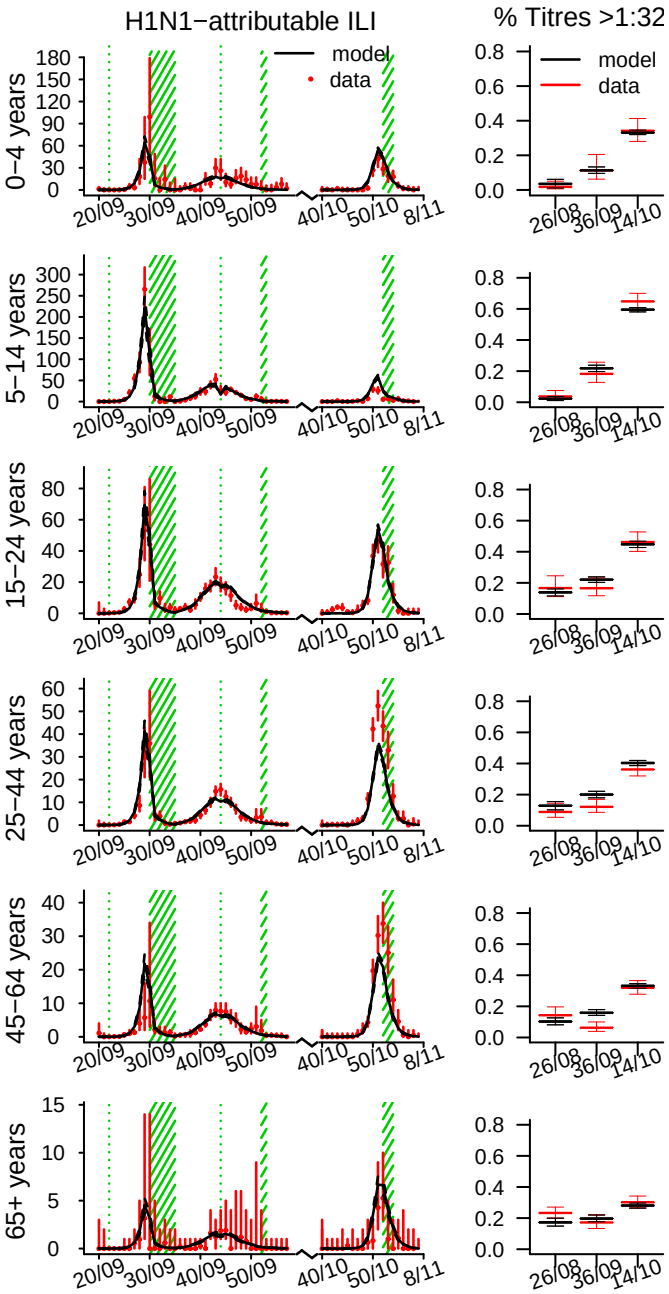


Figure S6: As Figure S4, but obtained using model variant 3.

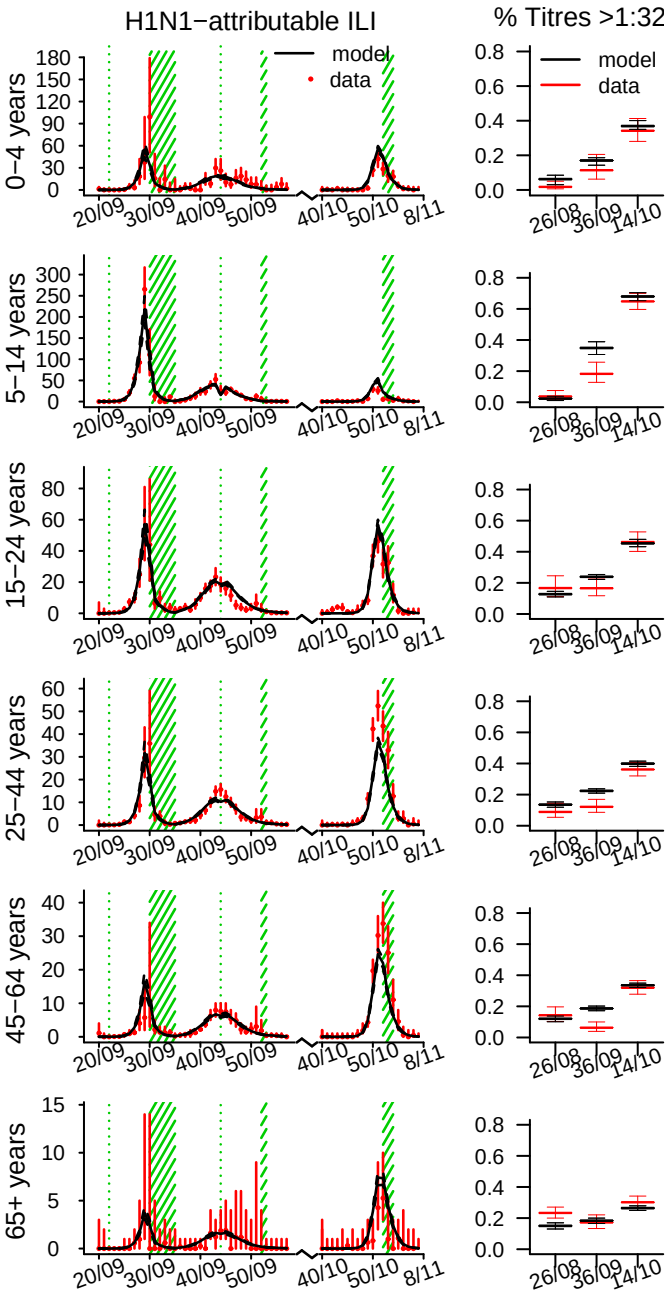


Figure S7: As Figure S4, but obtained using model variant 4.

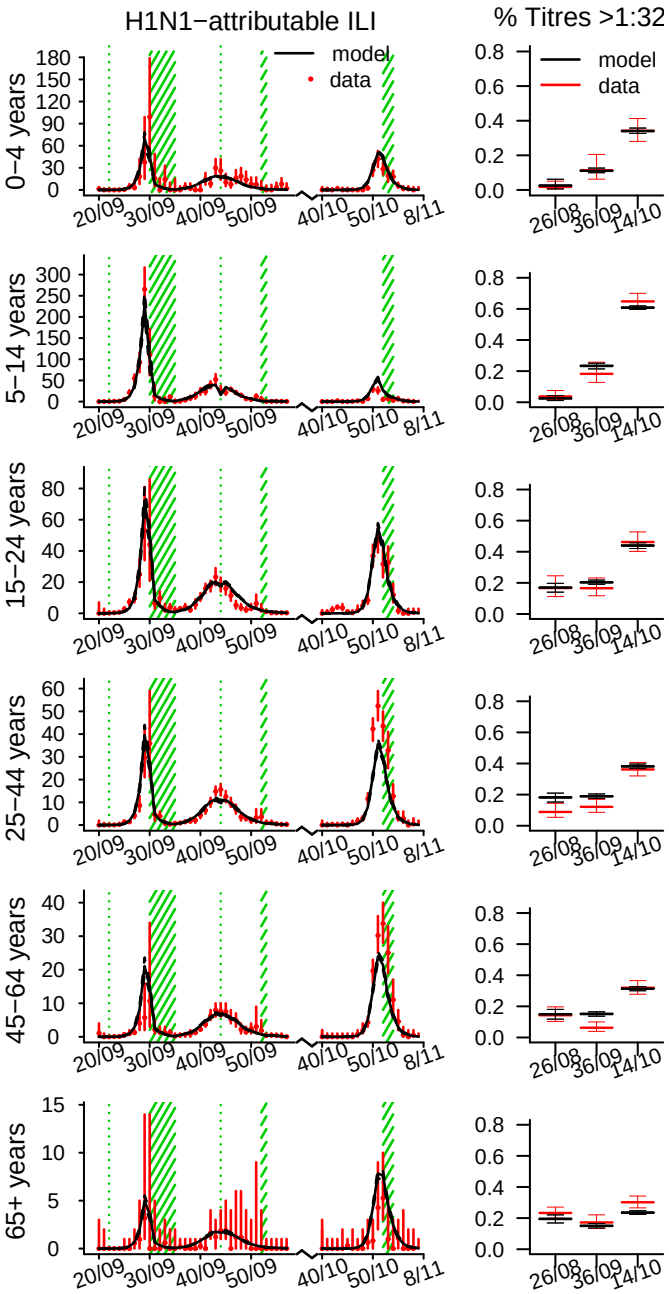


Figure S8: As Figure S4, but obtained using model variant 5.

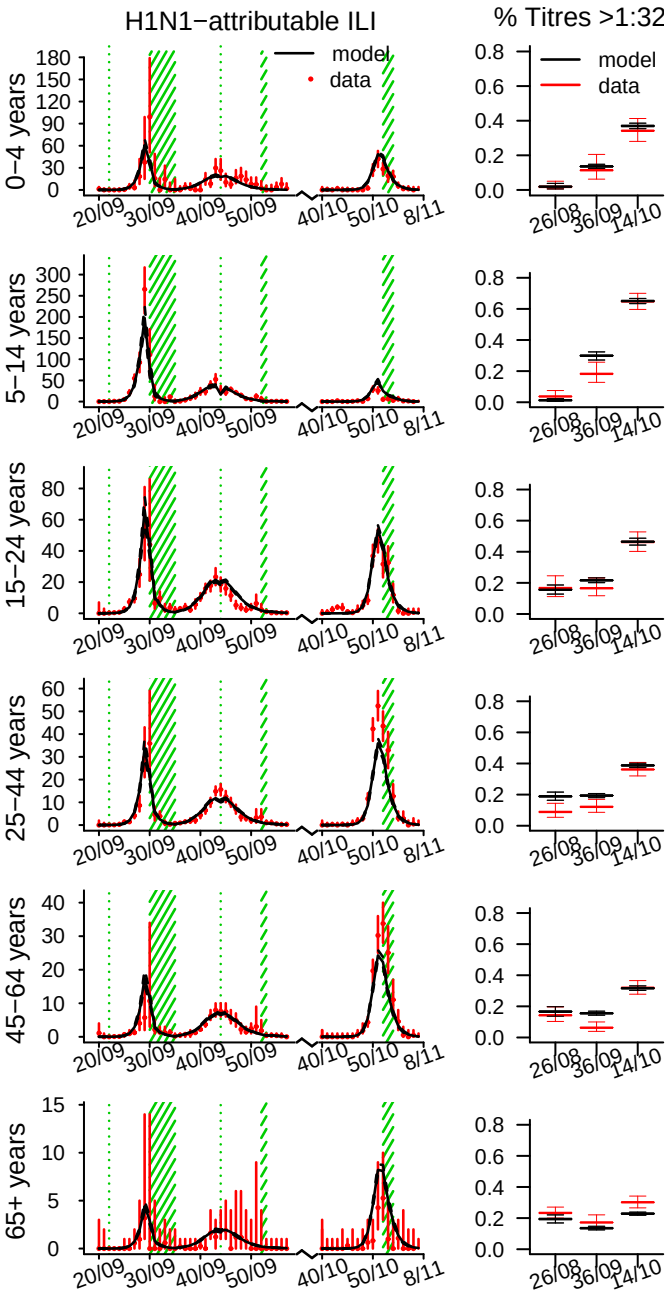


Figure S9: As Figure S4, but obtained using model variant 6.

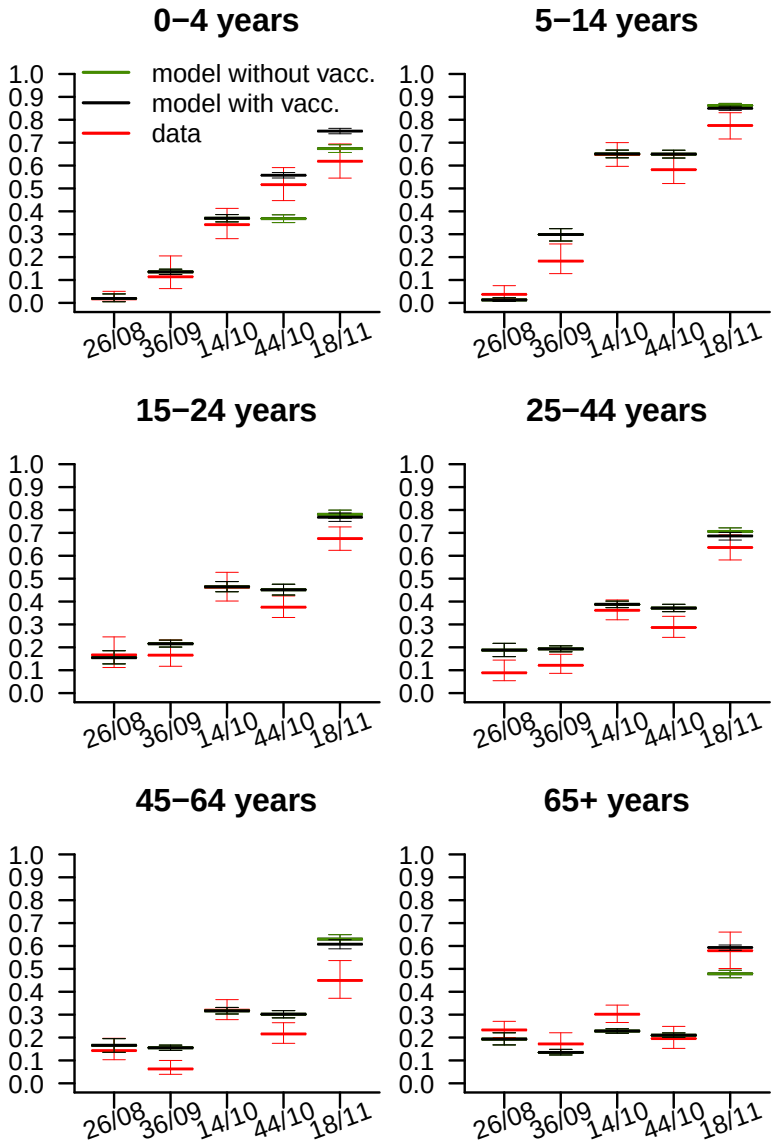


Figure S10: Plot of the average and 95% CI of the proportion of samples testing seropositive (HI tire $> 1:32$) by age group. Comparison of data (red) and model predictions obtained with model 6 without vaccination (green) and including vaccination (black). The first three data points of each panel representing pre-pandemic baseline (week 26/08), post-first wave (week 36/09) and post-second wave (week 14/10) serology have been fitted; the last two data points representing pre-third wave (week 44/10) and post-third wave (week 18/11) serology are predictions.

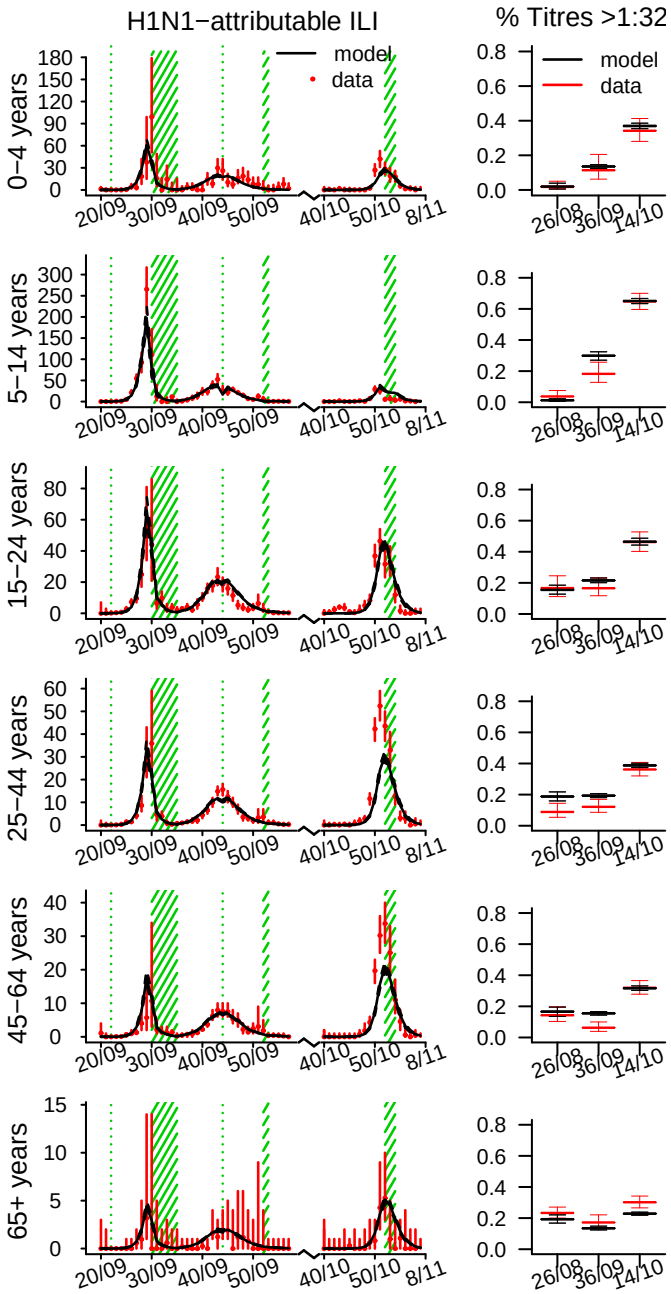


Figure S11: As Figure S4, but obtained using model variant 6 with vaccination.

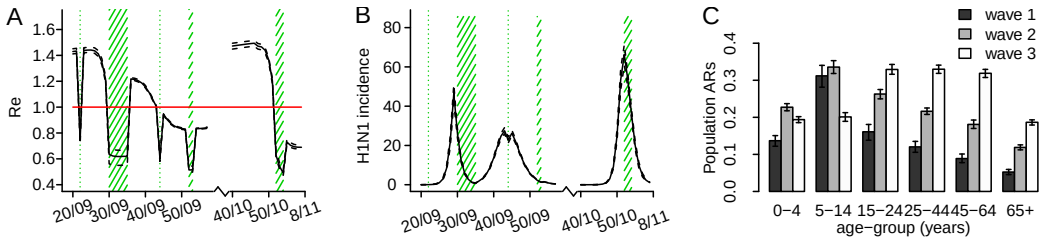


Figure S12: Estimates of (A) effective reproduction number in time (the solid line represents the mean, the dashed lines represent the 95% CrI, the red line is drawn at the threshold value of 1 and the green shaded areas represent holiday weeks); (B) incidence of H1N1 cases in the population (x 1000) in time (the solid line represents the mean, the dashed lines represent the 95% CrI and the green shaded areas represent holiday weeks); (C) attack rates by age group and wave (mean and 95% CrI) obtained with model 6 with vaccination.

2.2 Models with decay of homologous immunity: fit to the 2009-2011 ILI consultation and virological data and to the 2009-2010 serological data

In section 2.1 we present the results obtained with model variants 1-6 under the assumption that homologous immunity does not wane in time.
Here, we relax this assumption and for each model 1-6, we define the respective variant (denoted respectively 1D-6D) that allows homologous immunity to decay at a rate δ_2 to be estimated from the data. The flow diagram of the transmission model with decay of homologous immunity is given in Figure S13. Models 1D-6D are defined in Table S3.

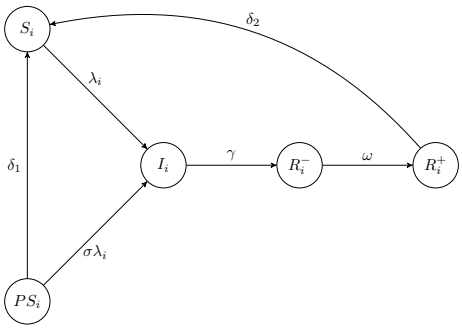


Figure S13: Flow diagram describing the transmission model. We assume that at the start of the pandemic individuals are either fully susceptible (S) to infection by H1N1pdm09 or have some degree of prior immunity against infection by H1N1pdm09, in which case they are classified as partially susceptible (PS). After infection, individuals move to the infectious state (I) and eventually recover. We model the biological delay to seroconversion by splitting the class of recovered individuals into two sub-classes: R^- denotes the recovered and fully immune testing seronegative, R^+ denotes the recovered and fully immune testing seropositive. Parameter σ represents the susceptibility of class PS, λ represents the force of infection, γ represents the recovery rate, ω represents the rate of seroconversion, δ_1 represents the rate of decay of prior immunity and δ_2 represents the rate of decay of homologous immunity. Subscript i denotes the age-class ($i = 1, \ldots, 6$).

Table S3: Description of model variants 1D-6D and relative assumptions on virus transmissibility and decay of prior and homologous immunity

Model	Transmissibility	Decay of prior immunity	Decay of homologous immunity
1D	same for all waves	no decay	estimated
2D	same for all waves	estimated	estimated
3D	different for wave 3	no decay	estimated
4D	different for each wave	no decay	estimated
5D	different for wave 3	estimated	estimated
6D	different for each wave	estimated	estimated

In Table S4 we present the estimates obtained with models 1D-6D. Figures S14-S19 show the fit to the H1N1-attributable ILI and serological data obtained with models 1D-6D.

Table S4: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters for model variants 1D-6D

Model	1D	2D	3D	4D	5D	6D
DIC	10275.5	9840.3	8335.9	8262.5	8231.1	8163.8
Log-	-5139	-4930	-4170	-4151	-4130	-4126
like.	(-5146, -5133)	(-4937, -4925)	(-4178, -4164)	(-4158, -4145)	(-4139, -4124)	(-4134, -4120)
R_e	1.45	1.37	1.45	1.46	1.42	1.44
	(1.43, 1.47)	(1.35, 1.39)	(1.43, 1.47)	(1.44, 1.48)	(1.40, 1.44)	(1.41, 1.46)
p_1	0.122	0.095	0.051	0.057	0.051	0.054
	(0.116, 0.128)	(0.093, 0.098)	(0.048, 0.054)	(0.054, 0.061)	(0.049, 0.052)	(0.052, 0.056)
p_2/p_1	-	-	-	1.16	-	1.12
	-	-	-	(1.12, 1.20)	-	(1.08, 1.15)
p_3/p_1	-	-	1.92	2.23	1.61	1.84
	-	-	(1.88, 1.95)	(2.14, 2.32)	(1.56, 1.66)	(1.77, 1.91)
$t_{1/2}(\delta_1)$	-	2.13	-	-	0.95	2.12
(years)	-	(2.09, 2.17)	-	-	(0.83, 1.06)	(1.95, 2.4)
$t_{1/2}(\delta_2)$	2.7	84.6	184.6	211.2	222.2	259.0
(years)	(2.6, 2.8)	(26.3, 3612.8)	(50.3, 7841.3)	(53.7, 11272.2)	(59.6, 8898.1)	(69.1, 11244.9)
ρ_1^4	0.0117	0.0161	0.0137	0.0083	0.0114	0.0082
	(0.0101, 0.0135)	(0.0141, 0.0180)	(0.0121, 0.0155)	(0.0070, 0.0100)	(0.0101, 0.0127)	(0.0070, 0.0096)
ρ_2^4/ρ_1^4	0.25	0.31	0.31	0.59	0.36	0.56
	(0.21, 0.29)	(0.26, 0.35)	(0.27, 0.35)	(0.49, 0.71)	(0.32, 0.40)	(0.46, 0.65)
ρ_3^4/ρ_1^4	0.47	0.46	0.23	0.41	0.31	0.45
	(0.41, 0.55)	(0.40, 0.53)	(0.19, 0.26)	(0.34, 0.49)	(0.27, 0.36)	(0.37, 0.53)
ξ	0.11	0.17	0.18	0.15	0.21	0.19
	(0.10, 0.12)	(0.16, 0.19)	(0.16, 0.21)	(0.13, 0.17)	(0.18, 0.24)	(0.16, 0.21)
I_1	1.7	1.8	1.6	2.9	2.4	3.6
	(1.0, 2.7)	(1.0, 3.1)	(1.0, 2.4)	(1.5, 5.3)	(1.4, 3.7)	(1.8, 60.2)
I_3	13.9	18.7	1.7	1.5	1.8	1.6
	(2.2, 32.5)	(2.0, 40.6)	(1.0, 3.1)	(1.0, 2.5)	(1.0, 3.4)	(1.0, 3.0)
σ	0.32	0.10	0.57	0.47	0.27	0.34
	(0.30, 0.34)	(0.09, 0.12)	(0.50, 0.63)	(0.43, 0.51)	(0.20, 0.35)	(0.28, 0.41)
ps_b^1	0.82	0.54	0.20	0.46	0.12	0.15
	(0.67, 0.94)	(0.47, 0.62)	(0.07, 0.38)	(0.30, 0.63)	(0.029, 0.29)	(0.05, 0.27)
ps_b^2	0.96	0.73	0.13	0.16	0.09	0.09
	(0.94, 0.99)	(0.70, 0.76)	(0.06, 0.22)	(0.07, 0.27)	(0.04, 0.16)	(0.04, 0.16)
ps_b^3	0.99	0.97	0.78	0.84	0.85	0.79
	(0.97, 1.00)	(0.93, 1.00)	(0.63, 0.94)	(0.73, 0.95)	(0.72, 0.97)	(0.68, 0.91)
ps_b^4	0.99	0.99	0.72	0.90	0.95	0.90
	(0.98, 1.00)	(0.97, 1.00)	(0.53, 0.90)	(0.76, 0.99)	(0.86, 1.00)	(0.76, 0.99)
ps_b^5	0.99	0.96	0.58	0.80	0.80	0.79
	(0.96, 1.00)	(0.91, 1.00)	(0.43, 0.74)	(0.65, 0.95)	(0.63, 0.96)	(0.64, 0.94)
ps_b^6	0.99	0.98	0.98	0.99	0.97	0.98
	(0.96, 1.00)	(0.94, 1.00)	(0.92, 1.00)	(0.95, 1.00)	(0.90, 1.00)	(0.93, 1.00)
ϵ_1	0.60	0.34	0.04	0.06	0.03	0.05
	(0.36, 0.81)	(0.05, 0.65)	(0.00, 0.14)	(0.00, 0.24)	(0.00, 0.11)	(0.00, 0.17)
ϵ_2	0.998	0.997	0.99	0.99	0.99	0.99
	(0.991, 0.999)	(0.991, 0.999)	(0.98, 1.00)	(0.98, 1.00)	(0.97, 1.00)	(0.97, 1.00)
ϵ_3	0.20	0.36	0.47	0.54	0.40	0.51
	(0.03, 0.38)	(0.20, 0.46)	(0.32, 0.59)	(0.40, 0.66)	(0.25, 0.53)	(0.36, 0.63)
ϵ_4	0.57	0.57	0.62	0.66	0.63	0.66
	(0.55, 0.59)	(0.54, 0.60)	(0.58, 0.66)	(0.61, 0.71)	(0.59, 0.66)	(0.61, 0.70)

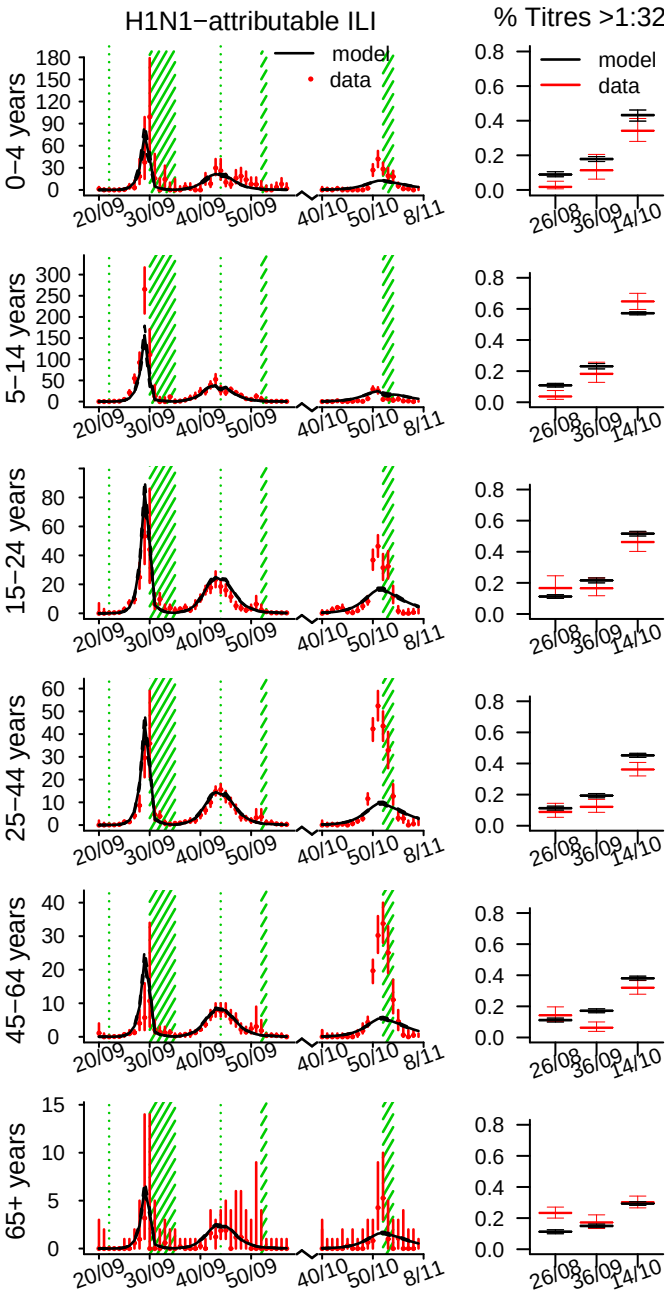


Figure S14: Age specific illness rates and seroprevalence estimated using model variant 1D. Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% CI around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1:32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09) and after the second wave (week 14/10).

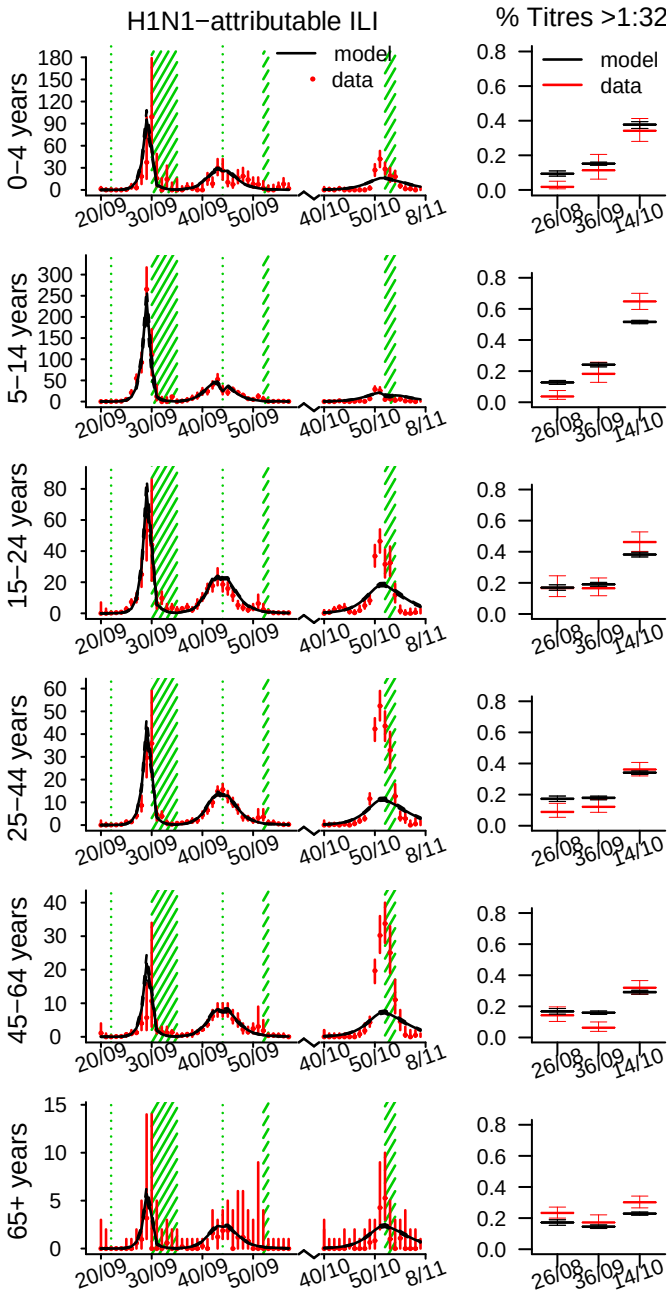


Figure S15: As Figure S14, but obtained using model variant 2D.

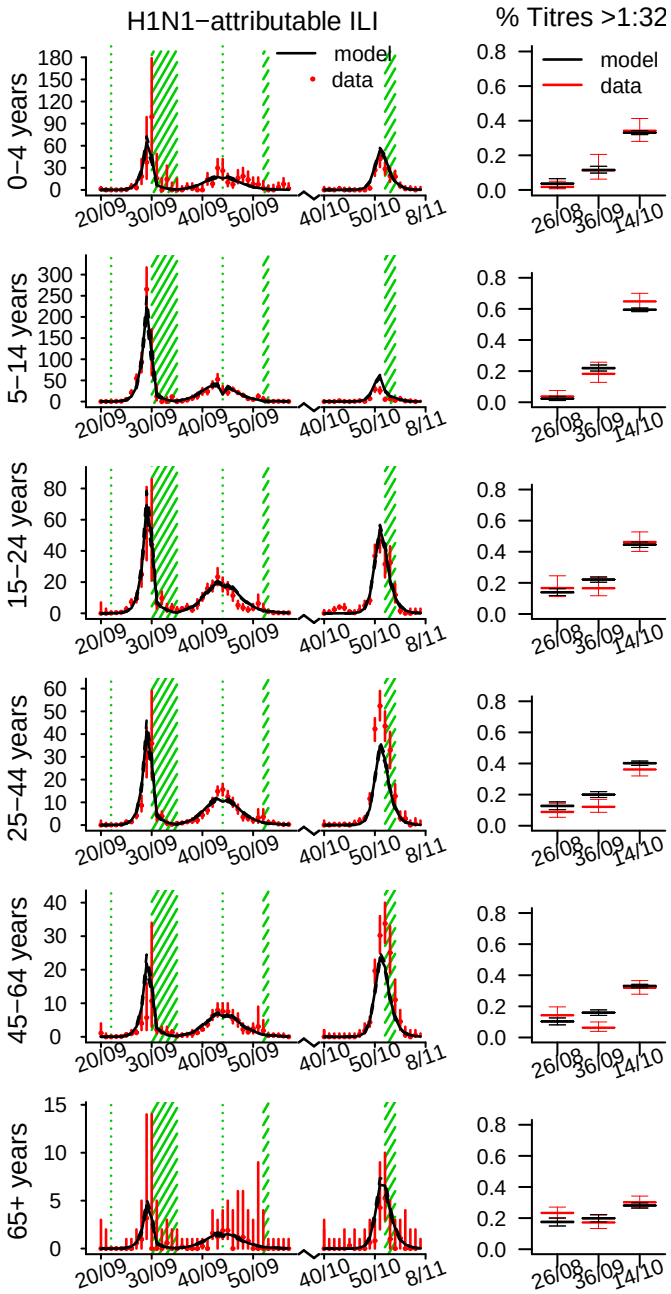


Figure S16: As Figure S14, but obtained using model variant 3D.

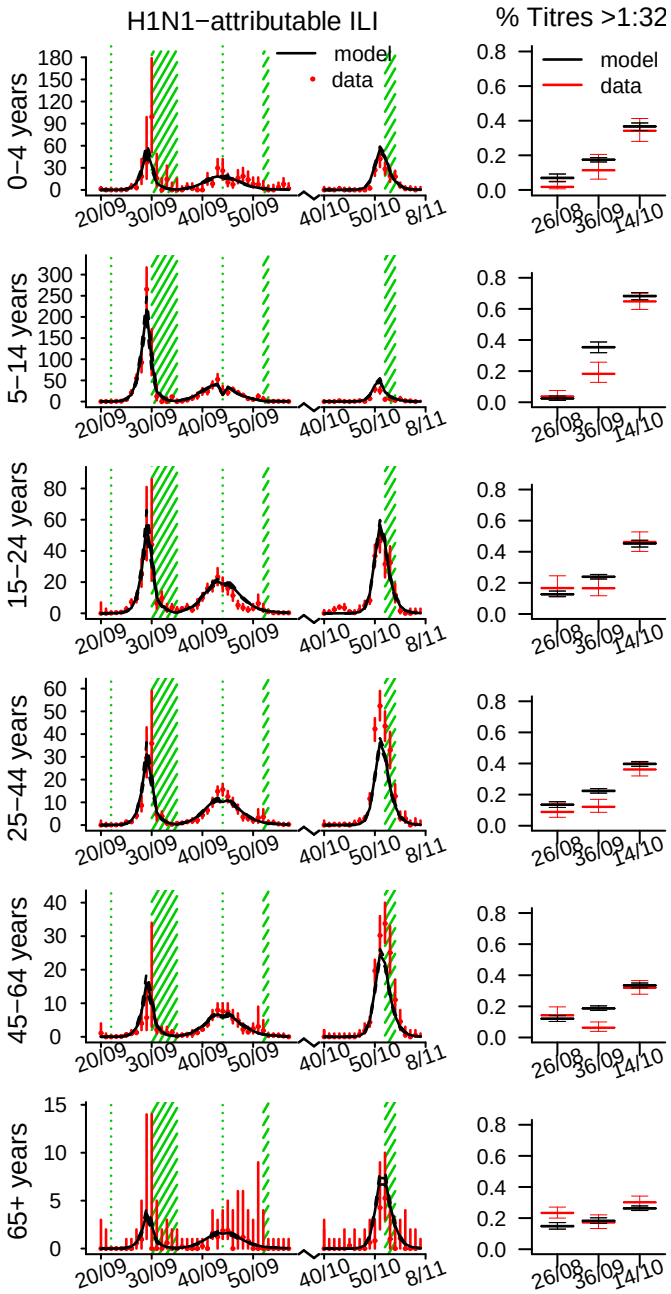


Figure S17: As Figure S14, but obtained using model variant 4D.

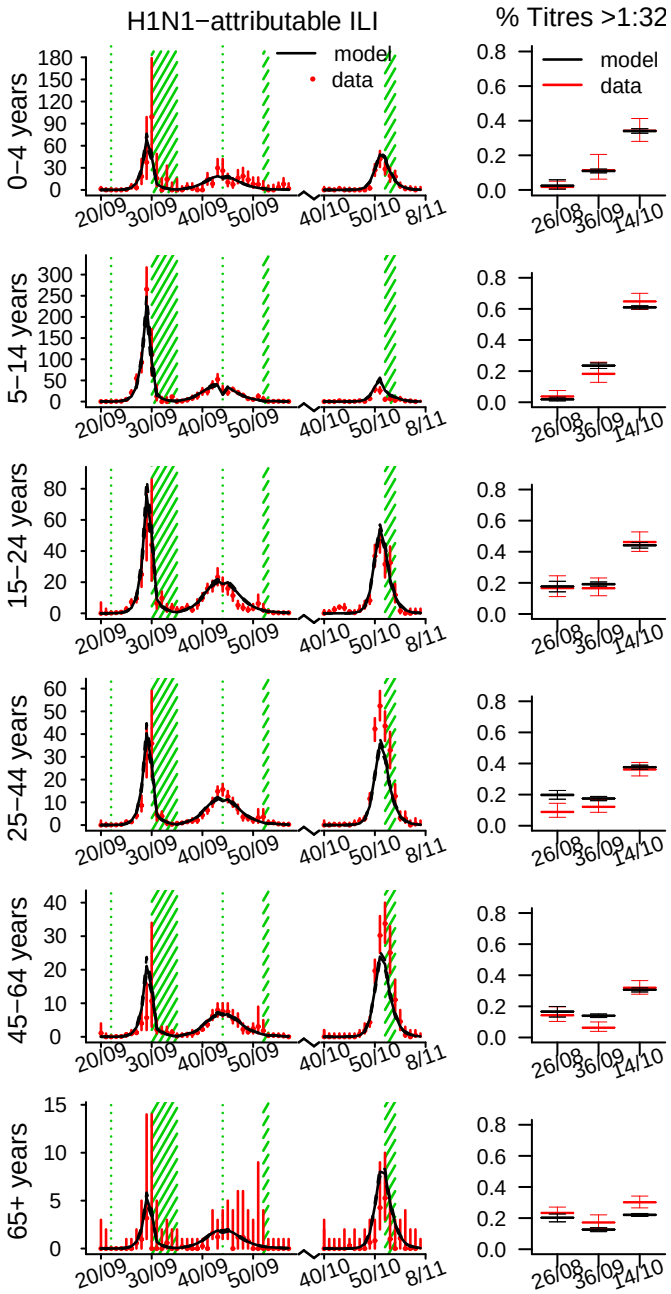


Figure S18: As Figure S14, but obtained using model variant 5D.

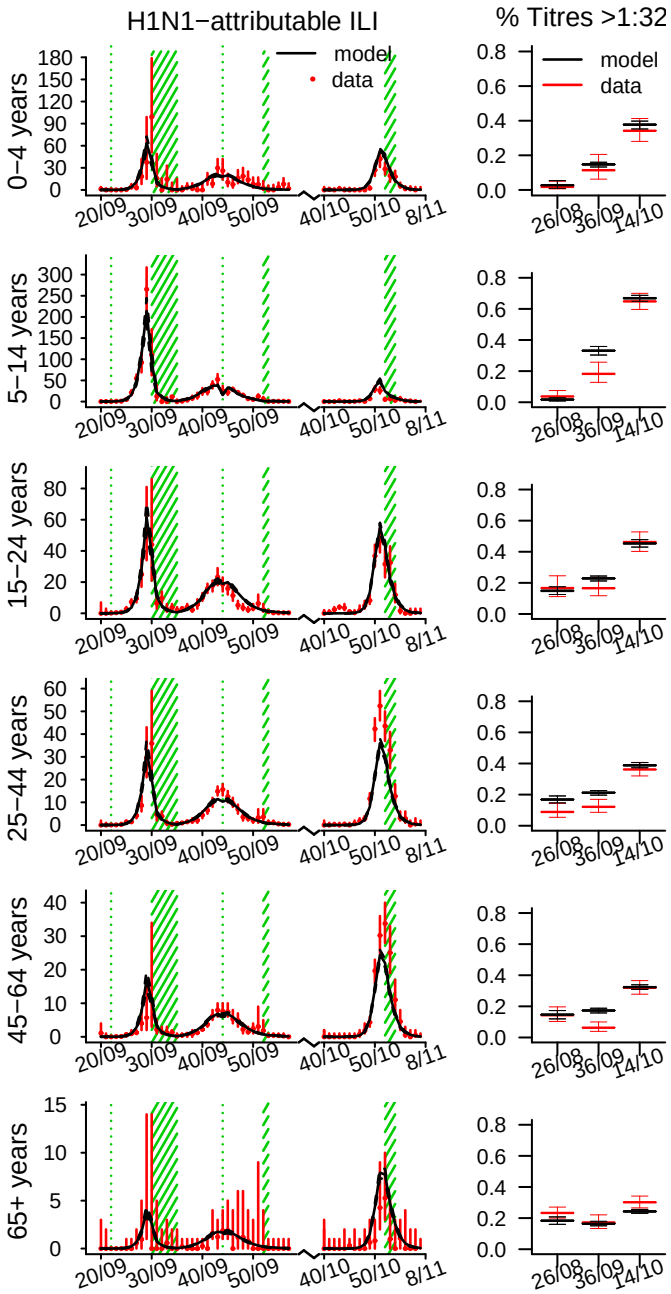


Figure S19: As Figure S14, but obtained using model variant 6D.

2.3 Fit of models 6 and 6D to the 2009-2011 ILI consultation, virological and serological data

Tables S5 and S6 show the posterior mean and 95% credible interval of the log-likelihood and parameters for models 6 and 6D respectively, in their variants with and without vaccination, when fitting the 2009-2011 ILI consultation, virological and serological data (i.e. including the third wave serological data in the fit, differently from what has been done in sections 2.1 and 2.2). Vaccination has been modelled as explained in the main text (see Vaccination in Materials and Methods). Figures S20 and S21 show the fit to the data obtained with model 6 without and with vaccination respectively. Figures S22 and S23 show the fit to the data obtained with model 6D without and with vaccination respectively.

Table S5: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters for model 6 with and without vaccination obtained by fitting the 2009-2011 ILI, virological and serological data (i.e. serological data on the third wave are included in the fit).

Model 6	without vaccination	with vaccination
Log-likelihood	-4203 (-4210, -4197)	-4176 (-4184, -4171)
R_e	1.40 (1.38, 1.43)	1.39 (1.37, 1.41)
p_1	0.052 (0.051, 0.053)	0.050 (0.049, 0.052)
p_2/p_1	1.06 (1.03, 1.09)	1.06 (1.04, 1.08)
p_3/p_1	1.59 (1.54, 1.64)	1.61 (1.56, 1.66)
$t_{1/2}(\delta_1)$ (years)	0.98 (0.95, 1.02)	0.89 (0.85, 0.93)
ρ_1^4	0.0094 (0.0079, 0.0107)	0.0095 (0.0083, 0.0107)
ρ_2^4/ρ_1^4	0.47 (0.40, 0.56)	0.47 (0.40, 0.53)
ρ_3^4/ρ_1^4	0.43 (0.36, 0.52)	0.45 (0.38, 0.52)
ξ	0.20 (0.17, 0.22)	0.21 (0.19, 0.24)
I_1	4.7 (2.4, 8.0)	5.7 (3.1, 8.8)
I_3	2.0 (1.0, 4.3)	2.0 (1.0, 4.4)
σ	0.17 (0.11, 0.21)	0.08 (0.01, 0.13)
ps_b^1	0.08 (0.02, 0.16)	0.09 (0.03, 0.20)
ps_b^2	0.08 (0.04, 0.13)	0.06 (0.03, 0.10)
ps_b^3	0.82 (0.72, 0.93)	0.74 (0.65, 0.83)
ps_b^4	0.93 (0.83, 1.00)	0.88 (0.76, 0.97)
ps_b^5	0.86 (0.72, 0.98)	0.79 (0.65, 0.93)
ps_b^6	0.97 (0.89, 1.00)	0.96 (0.87, 1.00)
ϵ_1	0.03 (0.00, 0.11)	0.03 (0.00, 0.09)
ϵ_2	0.98 (0.96, 1.00)	0.99 (0.96, 1.00)
ϵ_3	0.40 (0.22, 0.54)	0.47 (0.32, 0.59)
ϵ_4	0.66 (0.62, 0.70)	0.65 (0.61, 0.68)

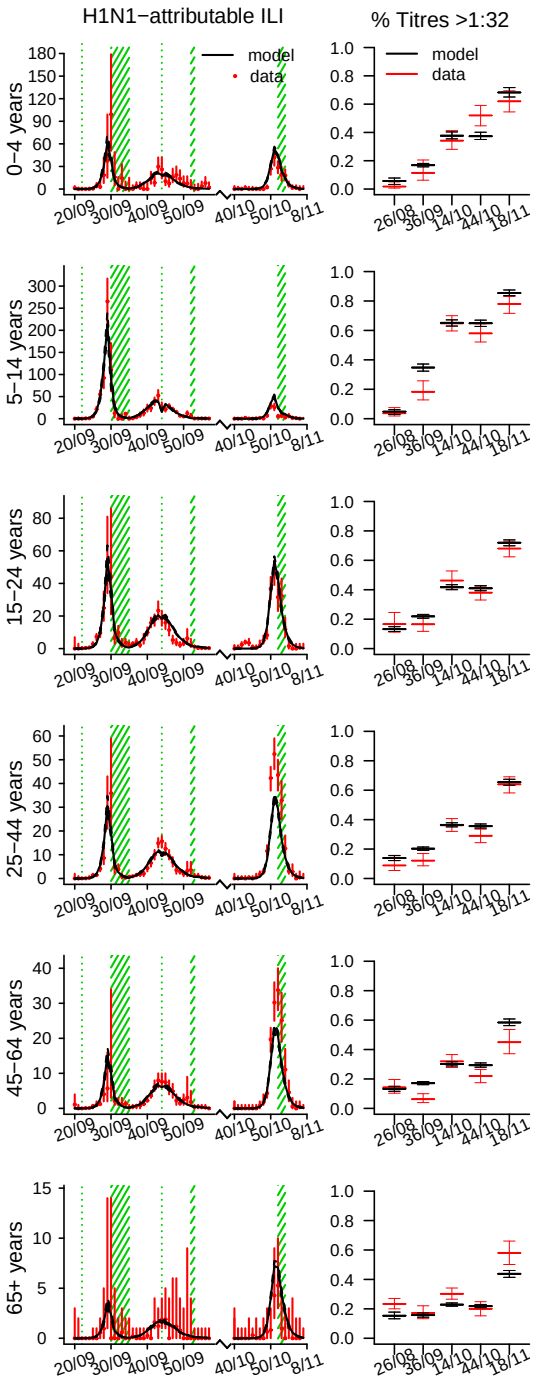


Figure S20: Age specific illness rates and seroprevalence estimated using model variant 6 without vaccination. Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% confidence interval around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1:32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09), after the second wave (week 14/10), before the third wave (44/10) and after the third wave (18/11).

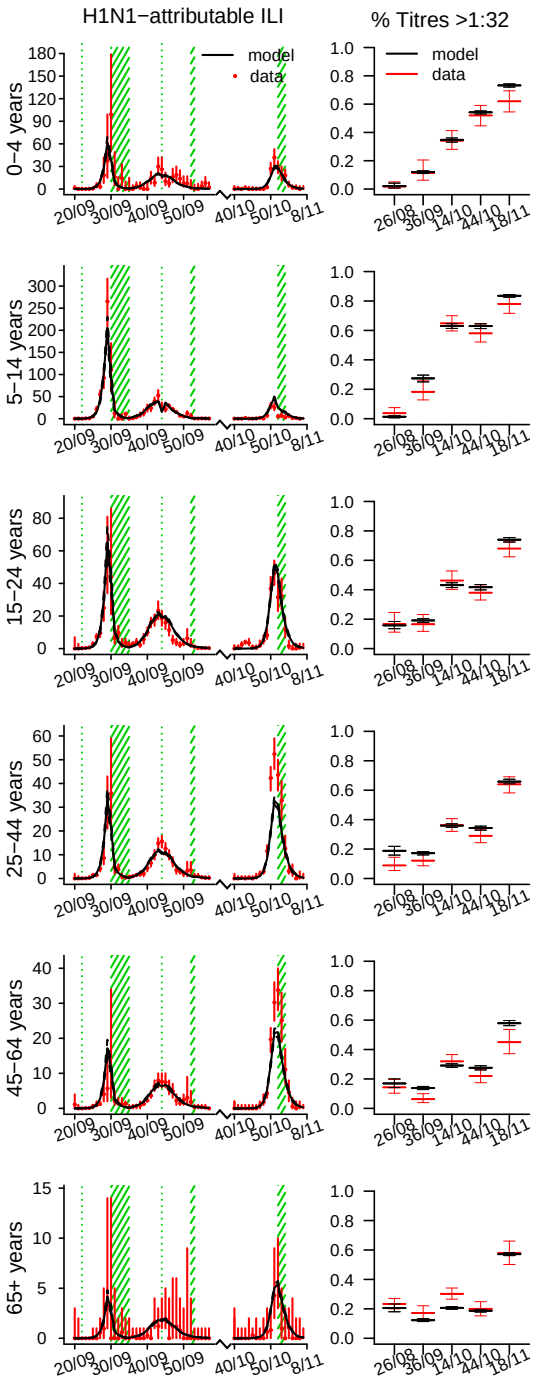


Figure S21: As Figure S20, but obtained using model variant 6 with vaccination.

Table S6: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters for model 6D with and without vaccination obtained by fitting the 2009-2011 ILI, virological and serological data (i.e. serological data on the third wave are included in the fit).

Model 6D	without vaccination	with vaccination
Log-likelihood	-4204 (-4211, -4199)	-4177 (-4186, -4171)
R_e	1.41 (1.39, 1.43)	1.39 (1.37, 1.42)
p_1	0.051 (0.050, 0.053)	0.050 (0.049, 0.052)
p_2/p_1	1.08 (1.05, 1.10)	1.04 (1.01, 1.07)
p_3/p_1	1.61 (1.56, 1.66)	1.61 (1.57, 1.66)
$t_{1/2}(\delta_1)$ (years)	1.00 (0.97, 1.06)	0.88 (0.83, 0.92)
$t_{1/2}(\delta_2)$ (years)	212.2 (59.6, 7447.7)	192.5 (52.9, 6291.6)
ρ_1^4	0.0087 (0.0076, 0.0101)	0.0102 (0.0086, 0.0117)
ρ_2^4/ρ_1^4	0.51 (0.43, 0.59)	0.42 (0.35, 0.49)
ρ_3^4/ρ_1^4	0.46 (0.39, 0.54)	0.41 (0.34, 0.49)
ξ	0.21 (0.18, 0.24)	0.21 (0.18, 0.24)
I_1	4.6 (2.5, 7.4)	5.1 (2.8, 7.9)
I_3	1.9 (1.0, 4.0)	1.8 (1.0, 3.7)
σ	0.12 (0.06, 0.17)	0.14 (0.08, 0.19)
ps_b^1	0.07 (0.02, 0.17)	0.11 (0.04, 0.19)
ps_b^2	0.06 (0.03, 0.11)	0.07 (0.03, 0.11)
ps_b^3	0.75 (0.65, 0.85)	0.79 (0.68, 0.90)
ps_b^4	0.86 (0.73, 0.97)	0.93 (0.82, 0.99)
ps_b^5	0.80 (0.65, 0.95)	0.83 (0.68, 0.97)
ps_b^6	0.96 (0.87, 1.00)	0.97 (0.89, 1.00)
ϵ_1	0.03 (0.00, 0.10)	0.03 (0.00, 0.09)
ϵ_2	0.99 (0.96, 1.00)	0.99 (0.96, 1.00)
ϵ_3	0.39 (0.24, 0.55)	0.48 (0.30, 0.59)
ϵ_4	0.67 (0.62, 0.70)	0.64 (0.60, 0.68)

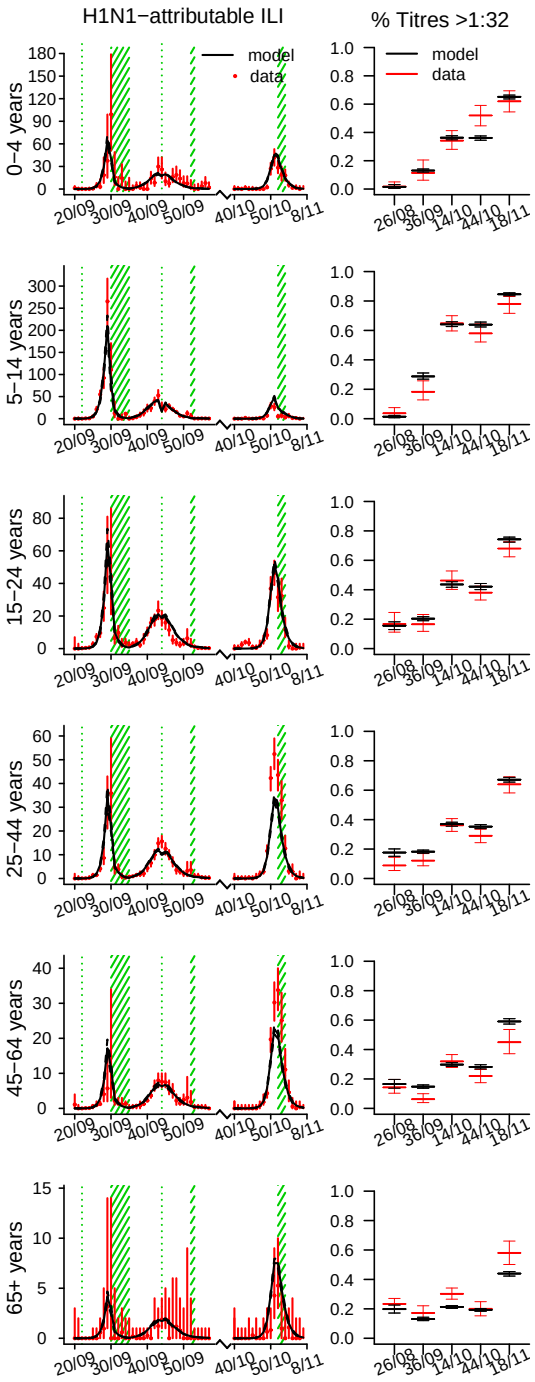


Figure S22: As Figure S20, but obtained using model variant 6D without vaccination.

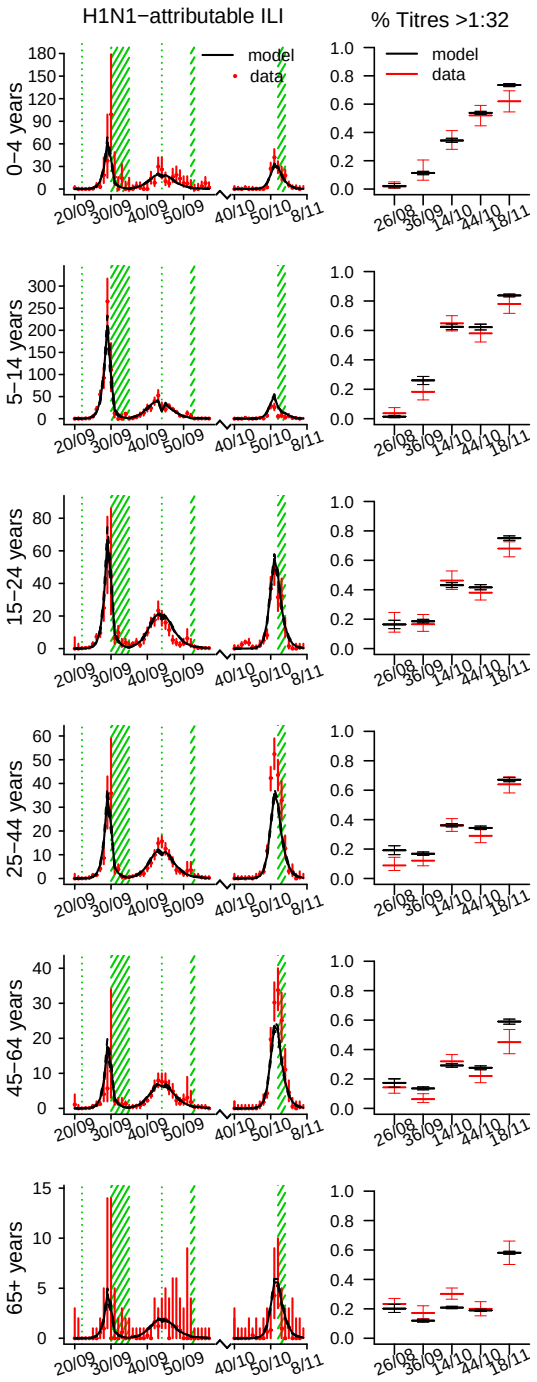


Figure S23: As Figure S20, but obtained using model variant 6D with vaccination.

3 Sensitivity Analysis

We performed sensitivity analysis on the reporting rates, on the representation of holidays and on the hypothesis that individuals seroconvert with certainty after infection. Sensitivity analysis on these assumptions was performed for model variant 6, as this was the variant with the highest mean posterior likelihood. As in sections 2.1-2.2, we fitted the 2009-2011 ILI consultation and virological data and the 2009-2010 serological data. The results are summarized in Tables S7 and S8.

3.1 Reporting rates

In the model presented in the main text we assumed that the the reporting rates vary across the age groups according to the propensity of the different age groups to visit the GP due to ILI observed in waves 1 and 2 (see section 1.5.1). Here we relax this assumption and estimate the reporting rate of under 15s in the third wave along with the other parameters of the model (i.e. we still assume that adults ($>$ 15 years) report ILI according to the observed propensities throughout the three waves). We further assume that $0-4$ and $5-14$ years old equally report ILI symptoms in the third wave (i.e. $\rho_3^1 = \rho_3^2$). Hence, the reporting rates to estimate from the data are ρ_3^1, ρ_1^4, ρ_2^4 and ρ_3^4. The parameters estimates obtained from the posterior distribution are summarized in Table S7 and Figure S24 shows the fit to the data obtained with this model.

3.2 Representation of holidays

We explore here a more parsimonious representation of holidays. For any holiday, we scale the contact rate of school-aged children with school-aged children ($5-14$ years) by a factor ϵ_5 and rescale all other contact rates by a different reduction factor ϵ_6. The parameters estimates are summarized in Table S7 and Figure S25 shows the fit to the data obtained with this model.

3.3 Seroconversion probability

In the model presented in the main text we assumed that after infection individuals seroconvert with certainty. Here we relax this hypothesis by assuming that infection by H1N1pdm09 virus does not necessary lead to seroconversion. The structure of the transmission model is as before, the only difference being in the definition of the infectious state where infected individuals eventually flow. As before, we assume that after infection individuals test seronegative (R^-) on average for 2 weeks and then move to state R* where the probability of testing seropositive is ν. The transmission model analyzed here consists of the following differential equations:

$$\begin{aligned}
\dot{S}_i &= \delta_1 P S_i - \lambda_i(t) S_i \\
\dot{PS}_i &= -\delta_1 P S_i - \sigma \lambda_i(t) P S_i \\
\dot{I}_i &= \lambda_i(t)(\sigma P S_i + S_i) - \gamma I_i \\
\dot{R^-}_i &= \gamma I_i - \omega R_i^- \\
\dot{R^*}_i &= \omega R_i^- .
\end{aligned} \tag{22}$$

The probability of testing seropositive (see eq. (17) for comparison) is now given by

$$w_t^i = \xi \frac{PS_i(t)}{N_i} + \nu \frac{R_i^*(t)}{N_i} \tag{23}$$

All other assumptions and the model parameterization are the same as in the main model (2). The parameters estimates are summarized in Table S7 and Figure S26 shows the fit to the data obtained with this model.

Table S7: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters obtained in the sensitivity analysis

		Sensitivity Analysis	
	Reporting	Holidays	Seropositivity
Log-likelihood	-4028 (-4035, -4022)	-4231 (-4237, -4226)	-4118 (-4126, -4112)
R_e	1.42 (1.39, 1.44)	1.37 (1.36, 1.38)	1.41 (1.39, 1.43)
p_1	0.0516 (0.0500, 0.0533)	0.0491 (0.0485, 0.0497)	0.0509 (0.0497, 0.0521)
p_2/p_1	1.05 (1.02, 1.08)	1.13 (1.10, 1.17)	1.09 (1.05, 1.13)
p_3/p_1	1.74 (1.68, 1.81)	1.74 (1.69, 1.80)	1.71 (1.64, 1.77)
$t_{1/2}(\delta_1)$	1.77 (1.55, 2.20)	0.95 (0.93, 0.96)	0.97 (0.93, 1.02)
ρ_1^4	0.0095 (0.0080, 0.0111)	0.0057 (0.0049, 0.0067)	0.0077 (0.0064, 0.0095)
ρ_2^4/ρ_1^4	0.48 (0.41, 0.56)	0.78 (0.66, 0.93)	0.55 (0.44, 0.66)
ρ_3^4/ρ_1^4	0.50 (0.42, 0.59)	0.80 (0.66, 0.97)	0.51 (0.40, 0.63)
ρ_3^1/ρ_3^4	0.81 (0.71, 0.92)	-	-
ξ	0.24 (0.21, 0.27)	0.19 (0.17, 0.22)	0.20 (0.17, 0.24)
ν	-	-	0.95 (0.89, 0.99)
I_1	4.0 (1.9, 7.0)	20.0 (14.6, 25.3)	4.6 (2.3, 7.7)
I_3	2.2 (1.1, 4.3)	1.2 (1.0, 2.0)	1.6 (1.0, 3.2)
σ	0.19 (0.12, 0.26)	0.22 (0.16, 0.27)	0.22 (0.14, 0.29)
ps_b^1	0.08 (0.02, 0.16)	0.09 (0.03, 0.17)	0.09 (0.03, 0.17)
ps_b^2	0.11 (0.06, 0.17)	0.03 (0.02, 0.06)	0.06 (0.03, 0.11)
ps_b^3	0.55 (0.46, 0.64)	0.78 (0.64, 0.91)	0.77 (0.63, 0.92)
ps_b^4	0.62 (0.50, 0.76)	0.96 (0.87, 1.00)	0.94 (0.82, 1.00)
ps_b^5	0.55 (0.43, 0.69)	0.91 (0.76, 0.99)	0.84 (0.65, 0.98)
ps_b^6	0.95 (0.86, 1.00)	0.97 (0.90, 1.00)	0.97 (0.90, 1.00)
ϵ_1	0.19 (0.01, 0.53)	-	0.04 (0.00, 0.12)
ϵ_2	0.99 (0.98, 1.00)	-	0.99 (0.97, 1.00)
ϵ_3	0.37 (0.16, 0.53)	-	0.43 (0.25, 0.57)
ϵ_4	0.66 (0.62, 0.70)	-	0.65 (0.61, 0.69)
ϵ_5	-	0.03 (0.00, 0.10)	-
ϵ_6	-	0.76 (0.75, 0.78)	-

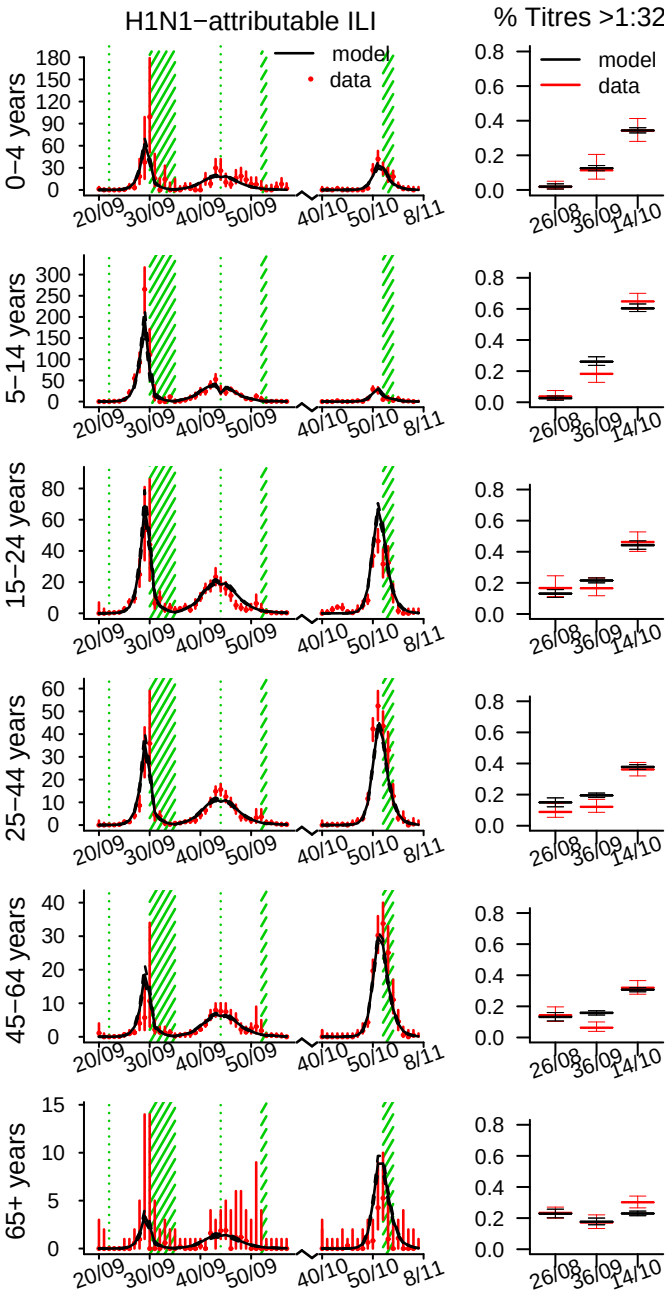


Figure S24: Age specific illness rates and seroprevalence estimated using model variant 6 with a less constrained parameterization of the reporting rates, as explained in section 3.1. Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% confidence interval around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1 : 32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09) and after the second wave (week 14/10).

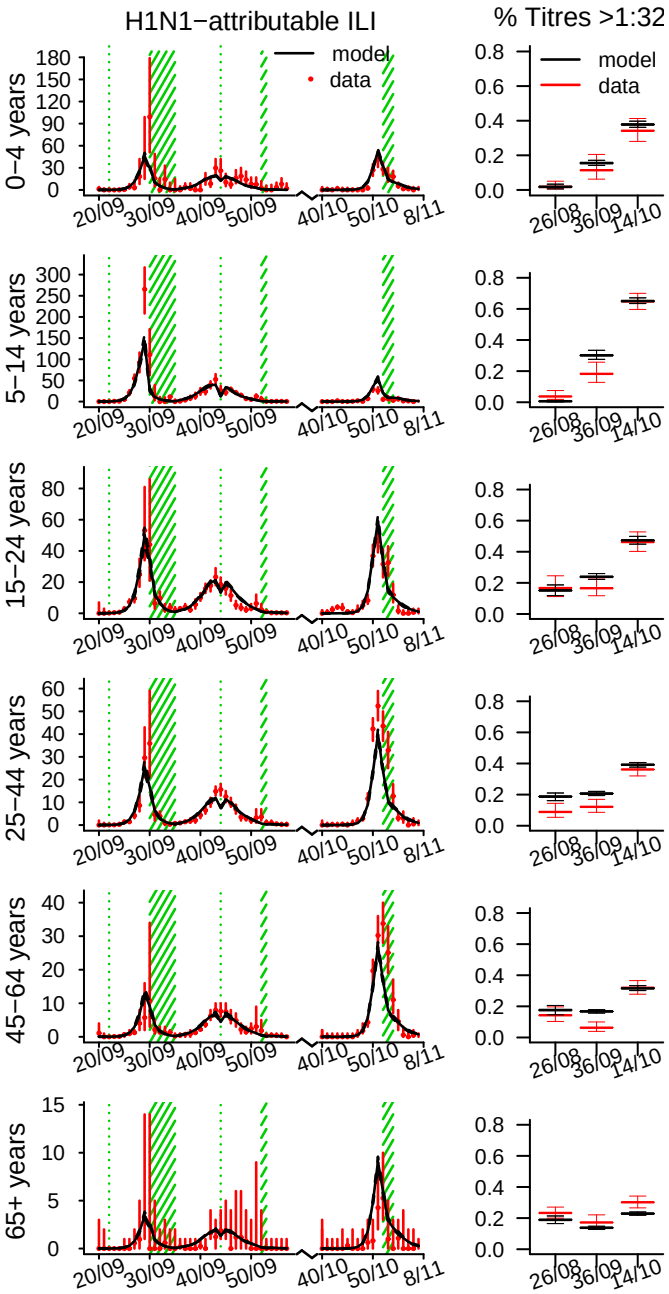


Figure S25: As Figure S24, but obtained using model variant 6 with a parsimonious representation of holidays, as explained in section 3.2.

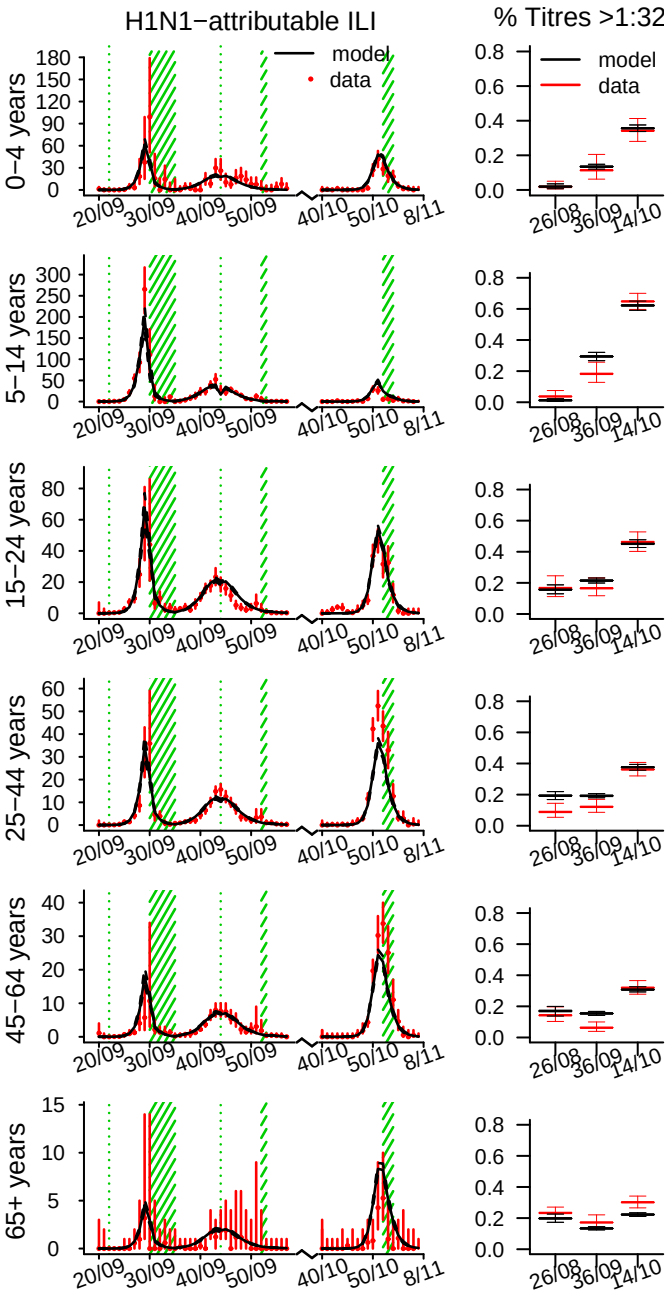


Figure S26: As Figure S24, but obtained using model variant 6 with the estimation of the probability of testing seropositive for infected and seroconverted individuals, as explained in section 3.3.

3.4 Assumed age distribution of seeded infections

In the model presented in the main text we assume that the initial numbers of cases I_1, I_3 seeding wave 1 and 3 respectively are distributed across the age classes proportionally to vector v =(5%, 40%, 30%, 10%, 10%, 5%). Here we test alternative distributions of the number of cases seeding infection across the age groups, according to vectors w_1 =(30%, 5%, 5%, 10%, 20%, 30%), w_2 =(16.7%, 16.7%, 16.6%, 16.6%, 16.7%,16.7%) and more skewed distributions, according to vectors w_3 =(0, 0, 0, 0, 50%,50%), w_4 =(50%,50%, 0, 0, 0, 0) and w_5 =(100%,0, 0, 0, 0, 0). We also tested the extreme case where the number of cases I_1 seeding wave 1 is distributed proportionally to vector w_3 and the number of cases I_3 seeding wave 3 is distributed across the age-classes proportionally to vector w_4, i.e. the distribution of the initial cases seeding infection in waves 1 and 3 is opposite to the observed distribution of cases in the respective waves (this case is denoted $w_3\&w_4$). The results are given in Table S8.

Table S8: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters obtained in the sensitivity analysis for alternative distributions across the age-classes of the number of cases seeding infection in waves 1 and 3

	w_1	w_2	w_3	w_4	w_5	w_3&w_4
Log-like.	-4117 (-4125, -4112)	-4118 (-4126, -4113)	-4118 (-4125,-4112)	-4117 (-4125, -4112)	-4118 (-4125, -4112)	-4118 (-4125, -4112)
R_e	1.42 (1.40, 1.44)	1.41 (1.39, 1.42)	1.40 (1.39, 1.43)	1.42 (1.40, 1.44)	1.41 (1.39, 1.43)	1.41 (1.39, 1.44)
p_1	0.051 (0.050, 0.053)	0.051 (0.049, 0.052)	0.051 (0.049, 0.052)	0.051 (0.050, 0.053)	0.051 (0.049, 0.052)	0.051 (0.049,0.052)
p_2/p_1	1.09 (1.06, 1.11)	1.06 (1.04, 1.09)	1.08 (1.05, 1.10)	1.09 (1.06, 1.12)	1.07 (1.05, 1.10)	1.07 (1.05, 1.10)
p_3/p_1	1.72 (1.66, 1.78)	1.68 (1.62, 1.74)	1.70 (1.65, 1.76)	1.71 (1.65, 1.76)	1.68 (1.63, 1.74)	1.69 (1.63, 1.74)
$t_{1/2}(\delta_1)$ (years)	1.20 (1.13, 1.33)	0.95 (0.90, 1.03)	1.03 (0.97, 1.08)	1.13 (1.09, 1.21)	0.93 (0.87, 1.00)	1.02 (0.91, 1.11)
ρ_1^4	0.0083 (0.0071, 0.0096)	0.0086 (0.0074, 0.0098)	0.0082 (0.0070, 0.0095)	0.0081 (0.0071, 0.0093)	0.0081 (0.0070, 0.0094)	0.0085 (0.0075, 0.0097)
ρ_2^4/ρ_1^4	0.52 (0.45, 0.61)	0.49 (0.42, 0.57)	0.52 (0.44, 0.60)	0.53 (0.46, 0.62)	0.52 (0.44, 0.61)	0.50 (0.43, 0.57)
ρ_3^4/ρ_1^4	0.45 (0.38, 0.54)	0.44 (0.37, 0.52)	0.46 (0.39, 0.55)	0.47 (0.40, 0.56)	0.47 (0.40, 0.56)	0.44 (0.38, 0.51)
ξ	0.20 (0.17, 0.23)	0.20 (0.18, 0.23)	0.20 (0.17, 0.23)	0.20 (0.17, 0.23)	0.20 (0.18, 0.23)	0.20 (0.18, 0.23)
I_1	9.2 (5.3, 13.5)	7.0 (3.9, 11.0)	17.3 (8.7, 26.3)	4.1 (2.1, 6.8)	11.0 (5.7, 17.3)	12.7 (6.3, 18.8)
I_3	2.0 (1.0, 4.5)	1.8 (1.0, 3.8)	2.2 (1.0, 5.0)	1.7 (1.0, 3.5)	2.1 (1.0, 4.7)	1.7 (1.0, 3.4)
σ	0.27 (0.22, 0.33)	0.24 (0.18, 0.30)	0.25 (0.18, 0.30)	0.24 (0.18, 0.30)	0.21 (0.17, 0.26)	0.24 (0.19, 0.28)
ps_b^1	0.09 (0.03, 0.19)	0.09 (0.03, 0.16)	0.09 (0.03, 0.19)	0.08 (0.03, 0.16)	0.08 (0.02, 0.18)	0.09 (0.02, 0.21)
ps_b^2	0.07 (0.03, 0.13)	0.07 (0.03, 0.12)	0.07 (0.03, 0.12)	0.07 (0.03, 0.11)	0.07 (0.03, 0.12)	0.07 (0.03, 0.12)
ps_b^3	0.79 (0.66, 0.93)	0.81 (0.69, 0.93)	0.80 (0.64, 0.94)	0.77 (0.64, 0.90)	0.79 (0.66, 0.93)	0.78 (0.66, 0.92)
ps_b^4	0.93 (0.82, 1.00)	0.96 (0.88, 1.00)	0.95 (0.82, 1.00)	0.92 (0.80, 1.00)	0.95 (0.85, 1.00)	0.94 (0.83, 1.00)
ps_b^5	0.82 (0.64, 0.98)	0.86 (0.71, 0.98)	0.85 (0.67, 0.99)	0.81 (0.65, 0.96)	0.85 (0.69, 0.98)	0.83 (0.65, 0.98)
ps_b^6	0.97 (0.91, 1.00)	0.97 (0.91, 1.00)	0.97 (0.90, 1.00)	0.97 (0.90, 1.00)	0.97 (0.91, 1.00)	0.97 (0.90, 1.00)
ϵ_1	0.04 (0.00, 0.12)	0.04 (0.00, 0.12)	0.04 (0.00, 0.12)	0.04 (0.00, 0.12)	0.04 (0.00, 0.12)	0.04 (0.00, 0.11)
ϵ_2	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)	0.99 (0.97, 1.00)
ϵ_3	0.46 (0.30, 0.59)	0.44 (0.28, 0.57)	0.45 (0.29, 0.58)	0.44 (0.28, 0.57)	0.42 (0.28, 0.55)	0.43 (0.27, 0.56)
ϵ_4	0.65 (0.61, 0.69)	0.64 (0.61, 0.68)	0.65 (0.61, 0.69)	0.65 (0.62, 0.69)	0.65 (0.62, 0.68)	0.65 (0.61, 0.68)

3.5 Out-of-fit predictions

In the manuscript we have already provided an important out-of-fit prediction, namely the prediction of the third wave serological data having fitted the model to the 2009-2011 ILI consultation and virological data and to the 2009-2010 serological data (see Fig. 4 in the manuscript and Fig. S10). Here we test how sensitive model results are to excluding other data, and the ability of the model to predict excluded data. We exclude each of the following in turn:

A. H1N1-attributable ILI of a single age group

B. H1N1-attributable ILI and post wave 1 and 2 seroprevalence in a single age group

C. post wave 1 and 2 seroprevalence of all age groups

D. wave 2 H1N1-attributable ILI of all age groups

The out-of-fit predictions for A and B are tested on 3 age groups (5-14 years, 15-24 years and 65+ years) using model variant 6. The results are presented in Tables S9-S11 and Figures S27-S34 show the model predictions.
We find that the model can successfully predict the H1N1-attributable ILI (A) of 15-24 year olds and 65+ year olds but fails to reproduce the high incidence observed at the peak of wave 1 in 5-14 year olds. Similarly, the model can successfully predict the H1N1-attributable ILI and the post wave 1 and 2 seroprevalence (B) of the 15-24 and 65+ groups but underestimates the incidence of H1N1-attributable ILI in the first wave and the post-second wave seroprevalence in the 5-14 group. In the absence of serological data other than pre-pandemic baseline seroprevalence (C), the model tends to overestimate the AR of wave 1 and underestimate the AR of wave 2. However, H1N1-attributable ILI of all age classes are fitted well. Finally, we find that in the absence of the wave 2 H1N1-attributable ILI of all age groups (D), the model cannot predict the H1N1-attributable ILI of wave 2 and in particular fails to reproduce the exponential growth rate of the ILI in the second wave (see Figure S34). This is unsurprising, as the serological data alone do not provide any information on the growth rate of the pandemic waves.

Table S9: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters obtained in the sensitivity analysis having left the H1N1-attributable ILI of the specified age groups out of the fit

A.	5-14 years	15-24 years	65+ years
Log-likelihood	-3431 (-3438, -3425)	-3342 (-3350, -3336)	-3591 (-3599, -3586)
R_e	1.38 (1.35, 1.41)	1.43 (1.40, 1.46)	1.41 (1.39, 1.43)
p_1	0.047 (0.044, 0.050)	0.052 (0.050, 0.054)	0.050 (0.049, 0.052)
p_2/p_1	1.02 (0.99, 1.05)	1.08 (1.04, 1.12)	1.08 (1.06, 1.11)
p_3/p_1	1.70 (1.62, 1.80)	1.71 (1.65, 1.78)	1.69 (1.63, 1.75)
$t_{1/2}(\delta_1)$	0.71 (0.57, 0.90)	1.13(0.96, 1.48)	1.00(0.91, 1.10)
ρ_1^4	0.0069 (0.0056, 0.0082)	0.0086 (0.0071, 0.0100)	0.0081 (0.0069, 0.0093)
ρ_2^4/ρ_1^4	0.59 (0.48, 0.72)	0.50 (0.40, 0.60)	0.53 (0.45, 0.62)
ρ_3^4/ρ_1^4	0.61 (0.49, 0.76)	0.46 (0.38, 0.57)	0.49 (0.41, 0.57)
ξ	0.23 (0.19, 0.27)	0.20 (0.17, 0.22)	0.20 (0.18, 0.23)
I_1	7.4 (3.1, 13.4)	2.7 (1.2, 5.5)	5.4 (3.0, 9.0)
I_3	2.0 (1.0, 4.0)	1.5 (1.0, 2.7)	1.8 (1.0, 3.6)
σ	0.16 (0.04, 0.34)	0.27 (0.21, 0.32)	0.23 (0.18, 0.28)
ps_b^1	0.07 (0.02, 0.16)	0.11 (0.03, 0.19)	0.09 (0.02, 0.20)
ps_b^2	0.16 (0.07, 0.29)	0.07 (0.03, 0.11)	0.06 (0.03, 0.10)
ps_b^3	0.54 (0.38, 0.70)	0.90 (0.74, 0.99)	0.78 (0.66, 0.91)
ps_b^4	0.78 (0.56, 0.98)	0.94 (0.82, 1.00)	0.95 (0.86, 1.00)
ps_b^5	0.69 (0.49, 0.91)	0.84 (0.67, 0.98)	0.84 (0.68, 0.98)
ps_b^6	0.95 (0.84, 1.00)	0.97 (0.92, 1.00)	0.96 (0.89, 1.00)
ϵ_1	0.83 (0.34, 0.99)	0.04 (0.00, 0.12)	0.04 (0.00, 0.12)
ϵ_2	0.99 (0.97, 1.00)	0.99 (0.96, 1.00)	0.99 (0.97, 1.00)
ϵ_3	0.10 (0.01, 0.29)	0.55 (0.41, 0.66)	0.46 (0.32, 0.59)
ϵ_4	0.67 (0.64, 0.70)	0.59 (0.54, 0.63)	0.65 (0.61, 0.68)

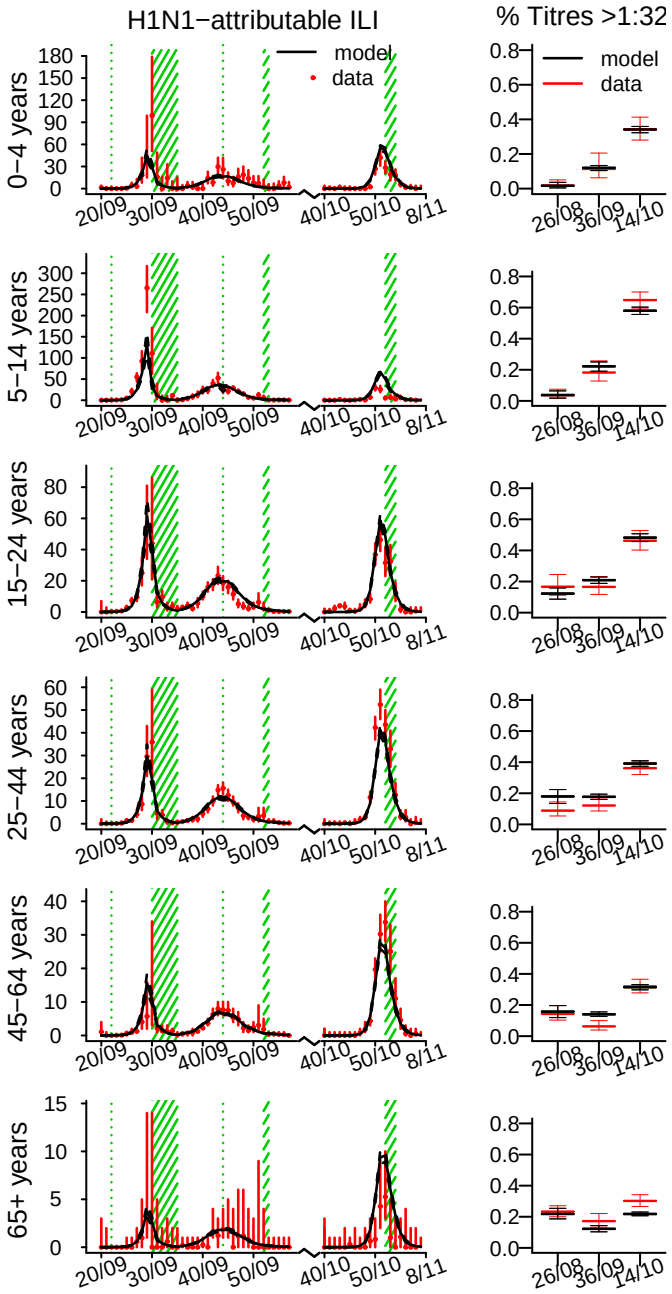


Figure S27: Age specific illness rates and seroprevalence predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data of all age groups except 5-14 years old and the 2009-2010 serological data, as explained in section 3.5 A. Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% confidence interval around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1:32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09) and after the second wave (week 14/10).

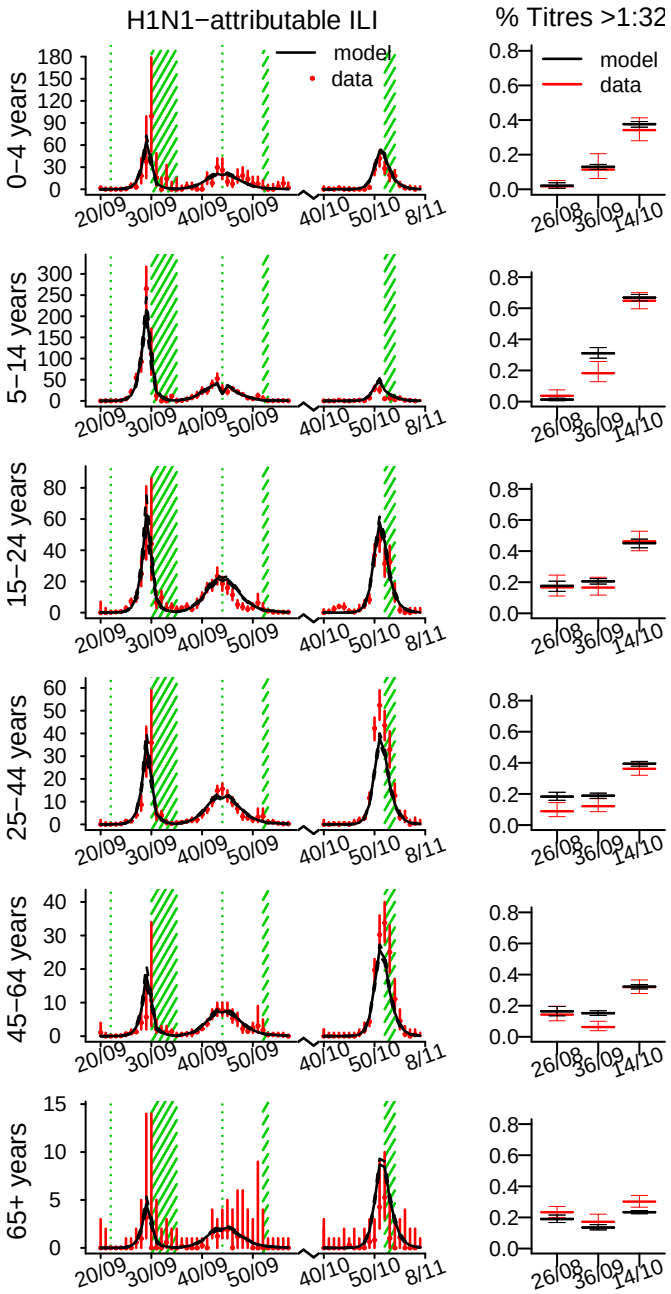


Figure S28: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data of all age groups except 15-24 years old and the 2009-2010 serological data, as explained in section 3.5 A.

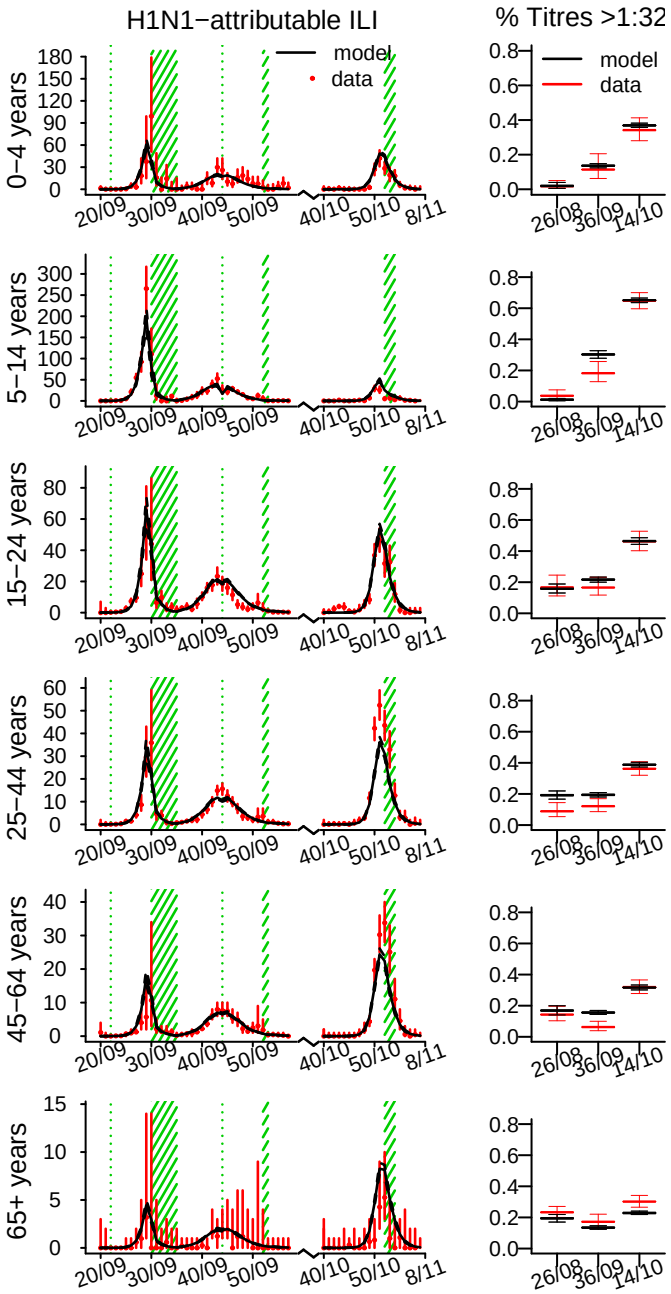


Figure S29: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data of all age groups except 65+ years old and the 2009-2010 serological data, as explained in section 3.5 A.

Table S10: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters obtained in the sensitivity analysis having left the H1N1-attributable ILI and serological data of the specified age groups out of the fit

B.	5-14 years	15-24 years	65+ years
Log-likelihood	-3421 (-3429, -3415)	-3338 (-3346, -3332)	-3574 (-3582, -3568)
R_e	1.40 (1.37, 1.43)	1.46 (1.44, 1.49)	1.39 (1.37, 1.42)
p_1	0.045 (0.042, 0.048)	0.056 (0.053, 0.059)	0.050 (0.048, 0.051)
p_2/p_1	0.98 (0.95, 1.02)	1.14 (1.10, 1.199)	1.06 (1.02, 1.11)
p_3/p_1	1.75 (1.65, 1.85)	1.86 (1.75, 1.98)	1.65 (1.59, 1.70)
$t_{1/2}(\delta_1)$	1.23 (0.67, 3.91)	2.23 (1.67, 3.17)	0.85 (0.78, 0.93)
ρ_1^4	0.0091 (0.0064, 0.0126)	0.0078 (0.0067, 0.0092)	0.0086 (0.0070, 0.0104)
ρ_2^4/ρ_1^4	0.44 (0.30, 0.62)	0.59 (0.48, 0.71)	0.49 (0.39, 0.61)
ρ_3^4/ρ_1^4	0.45 (0.30, 0.64)	0.51 (0.42, 0.61)	0.46 (0.35, 0.57)
ξ	0.29 (0.22, 0.43)	0.19 (0.16, 0.22)	0.19 (0.16, 0.21)
I_1	5.6 (2.3, 11.1)	1.9 (1.1, 3.4)	6.5 (3.1, 10.5)
I_3	2.5 (1.1, 5.2)	1.4 (1.0, 2.4)	1.8 (1.0, 3.7)
σ	0.37 (0.03, 0.68)	0.30 (0.22, 0.35)	0.19 (0.13, 0.25)
ps_b^1	0.07 (0.02, 0.15)	0.19 (0.07, 0.33)	0.09 (0.03, 0.16)
ps_b^2	0.14 (0.06, 0.26)	0.08 (0.03, 0.15)	0.07 (0.03, 0.11)
ps_b^3	0.41 (0.26, 0.59)	0.93 (0.75, 1.00)	0.79 (0.66, 0.93)
ps_b^4	0.41 (0.17, 0.72)	0.84 (0.68, 0.98)	0.95 (0.86, 1.00)
ps_b^5	0.43 (0.22, 0.67)	0.76 (0.60, 0.92)	0.89 (0.73, 0.99)
ps_b^6	0.85 (0.56, 0.99)	0.98 (0.92, 1.00)	0.97 (0.90, 1.00)
ϵ_1	0.84 (0.41, 0.99)	0.04 (0.00, 0.14)	0.03 (0.00, 0.11)
ϵ_2	0.99 (0.97, 1.00)	0.99 (0.96, 1.00)	0.99 (0.97, 1.00)
ϵ_3	0.10 (0.00, 0.29)	0.59 (0.46, 0.70)	0.45 (0.28, 0.59)
ϵ_4	0.67 (0.64, 0.70)	0.60 (0.54, 0.66)	0.65 (0.61, 0.69)

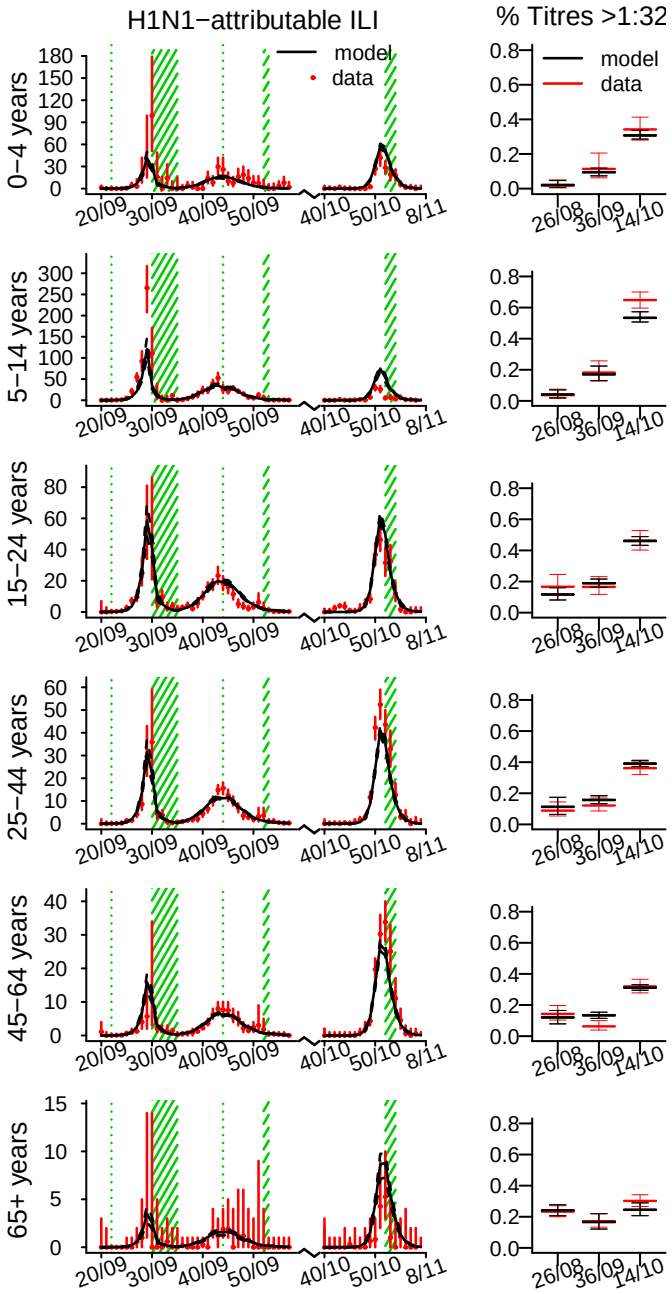


Figure S30: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data and the 2009-2010 serological data of all age groups except 5-14 years old, as explained in section 3.5 B.

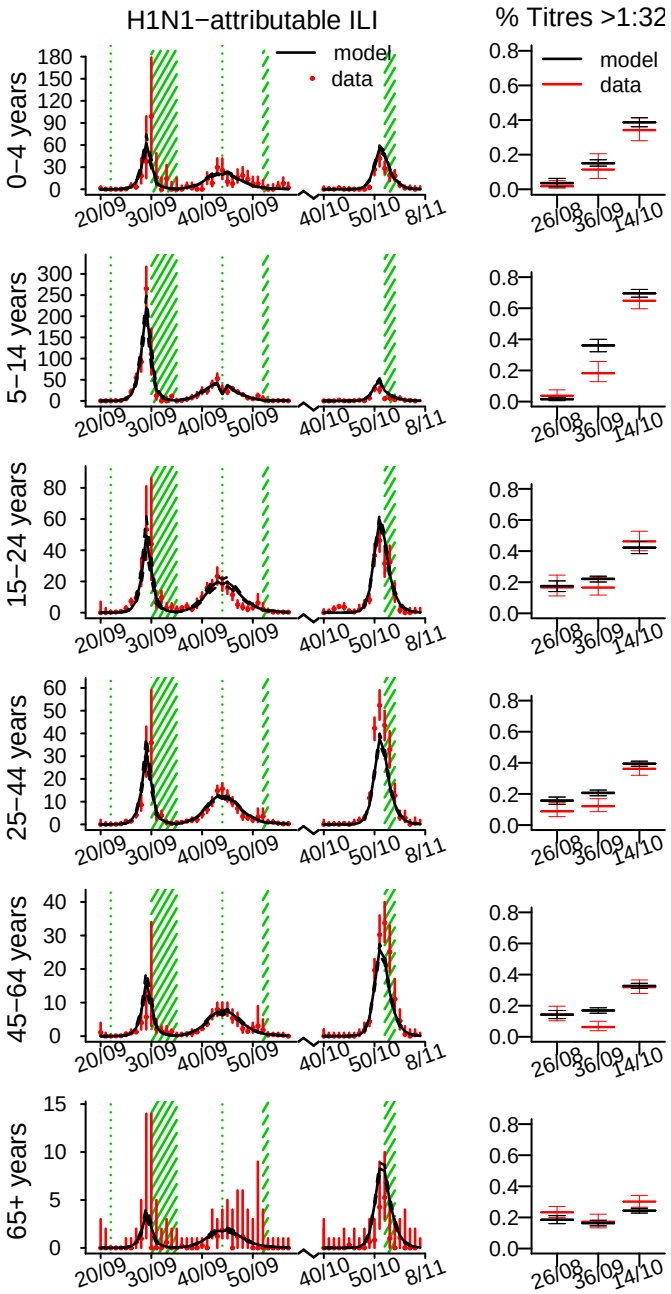


Figure S31: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data and the 2009-2010 serological data of all age groups except 15-24 years old, as explained in section 3.5 B.

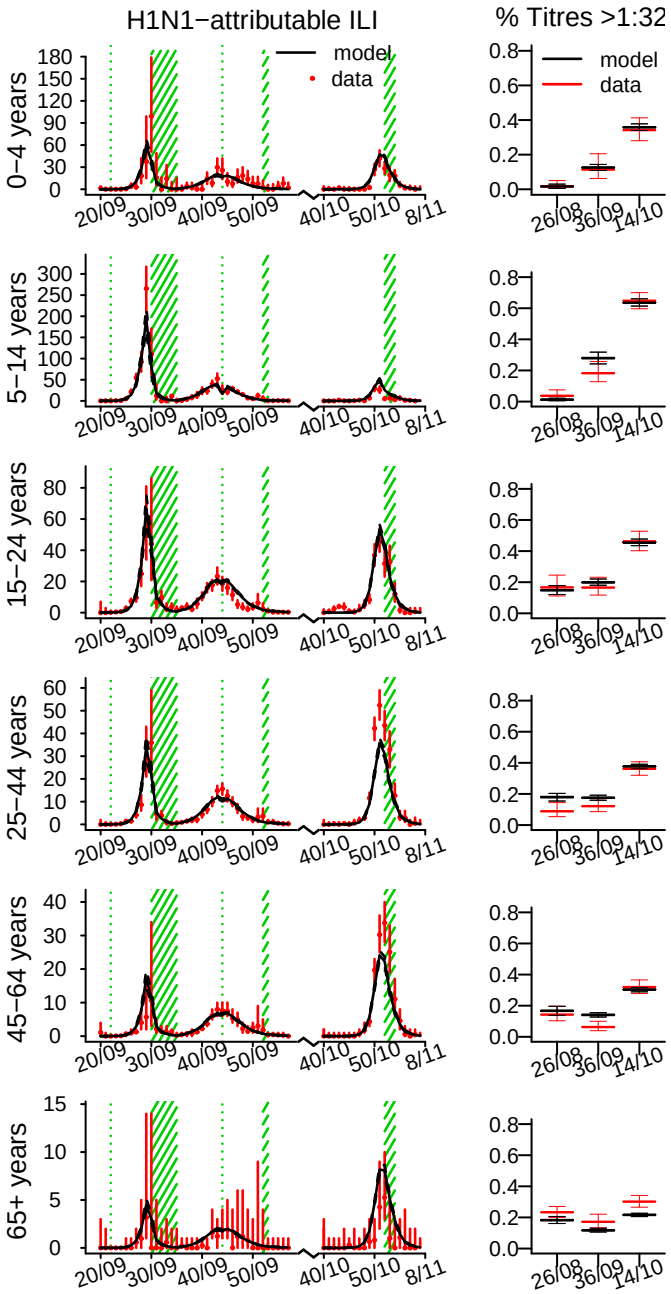


Figure S32: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data and the 2009-2010 serological data of all age groups except 65+ years old, as explained in section 3.5 B.

Table S11: Posterior mean and 95% CrI (within parentheses) of the log-likelihood and parameters obtained in the sensitivity analysis having left the wave 1 and 2 serological data of all age groups (C) and the wave 2 H1N1-attributable ILI of all age groups (D) out of the fit

	C	D
Log-likelihood	-4039 (-4047, -4033)	-2117 (-2126, -2112)
R_e	1.49 (1.46, 1.51)	1.36 (1.33, 1.39)
p_1	0.057 (0.054, 0.060)	0.054 (0.050, 0.058)
p_2/p_1	1.26 (1.22, 1.30)	1.28 (1.19, 1.38)
p_3/p_1	2.10 (2.02, 2.20)	1.98 (1.87, 2.12)
$t_{1/2}(\delta_1)$	3.22 (2.70, 3.96)	1.08 (0.83, 1.34)
ρ_1^4	0.0053 (0.0048, 0.0060)	0.0224 (0.0147, 0.0365)
ρ_2^4/ρ_1^4	0.93 (0.81, 1.00)	-
ρ_3^4/ρ_1^4	0.77 (0.65, 0.87)	0.23 (0.14, 0.33)
ξ	0.20 (0.17, 0.23)	0.20 (0.17, 0.22)
I_1	2.1 (1.1, 3.7)	5.3 (2.7, 9.3)
I_3	1.5 (1.0, 2.7)	2.9 (1.4, 5.4)
σ	0.37 (0.31, 0.43)	0.02 (0.00, 0.07)
ps_b^1	0.17 (0.08, 0.27)	0.21 (0.05, 0.38)
ps_b^2	0.10 (0.04, 0.17)	0.12 (0.06, 0.21)
ps_b^3	0.78 (0.68, 0.89)	0.67 (0.57, 0.79)
ps_b^4	0.94 (0.84, 1.00)	0.97 (0.89, 1.00)
ps_b^5	0.91 (0.78, 0.99)	0.90 (0.75, 0.99)
ps_b^6	0.98 (0.92, 1.00)	0.97 (0.90, 1.00)
ϵ_1	0.06 (0.00, 0.18)	0.04 (0.00, 0.15)
ϵ_2	0.99 (0.97, 1.00)	0.94 (0.89, 0.99)
ϵ_3	0.43 (0.24, 0.60)	0.38 (0.12, 0.62)
ϵ_4	0.72 (0.68, 0.76)	0.82 (0.68, 0.95)

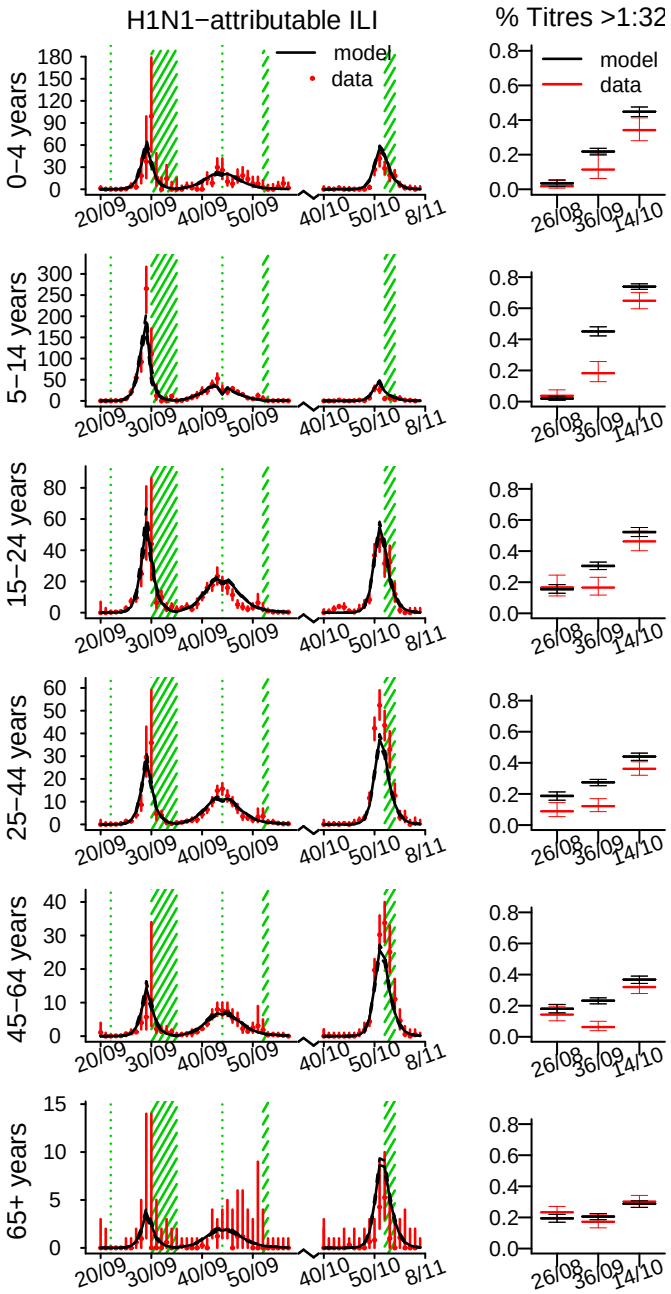


Figure S33: As Figure S27, but predicted using model variant 6 by fitting the 2009-2011 ILI consultation and virological data and baseline seroprevalence only, as explained in section 3.5 C.

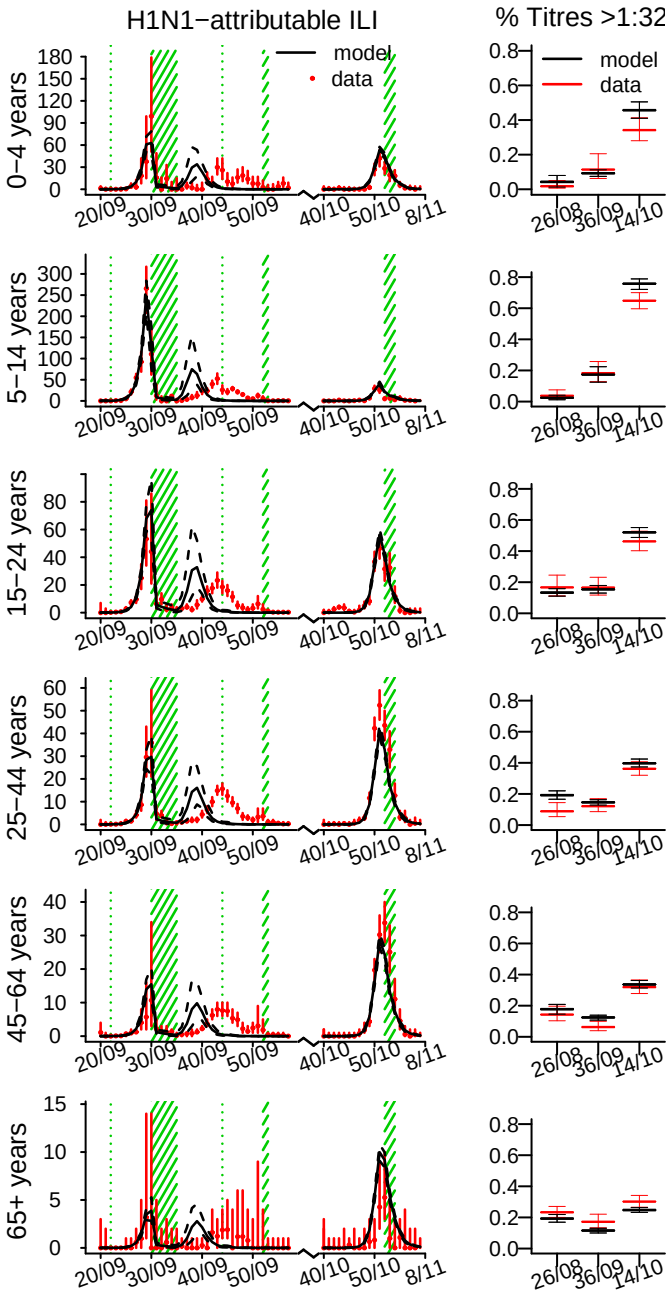


Figure S34: As Figure S27, but predicted using model variant 6 by fitting the ILI consultation and virological data of waves 1 and 3 and the 2009-2010 serological data, as explained in section 3.5 D. The reporting rates in wave 2 are assumed to be 10 times smaller than the reporting rates in wave 1 ($\rho_2^i = 0.1\rho_1^i$, $i = 1, \ldots, 6$).

3.6 Demographic stochasticity

From the literature, we know that the dynamics of a stochastic seasonally forced seasonal influenza model can differ substantially from the behaviour of the corresponding deterministic model [9, 10, 11, 12]. In order to obtain oscillations, seasonal influenza model need to assume an influx of susceptibles, for instance through immigration, vital demography or by waning of immunity (i.e. a SIRS model); in these scenarios sustained oscillations via coherence and dynamical resonance [9, 10] and stochastic amplification [11, 12] can occur. On the contrary, the classical closed population SIR model we adopt does not show the same stochastic sensitivity as can be seen in an open population seasonally forced epidemic model.

To verify this intuition, and in particular to test the role of stochasticity on the third wave, we undertook tests with a simulation study. We simulated the stochastic version of the deterministic model described in the manuscript in 52 million people, the population size of England. We randomly sampled 500 parameter sets from the posterior distribution of model variant 1 (the simplest, with no changes in transmissibility and no waning of immunity) and for each set of parameters we simulated 100 random realizations of the stochastic model. To avoid extinction, we assumed a small and constant external force of infection of order 10^{-5}/day. For the same 500 randomly drawn posterior parameter sets we also computed the solution of the deterministic model. Figure S35 shows the mean and 95%CI of H1N1 incidence (x 1000) obtained with the stochastic and deterministic models, with excellent agreement being seen between the deterministic solution and the average of the stochastic epidemics. Figure S35 also shows that stochastic effects alone do not reproduce the third wave.

Where the stochastic and deterministic models differ is in the variability around the mean behavior, which arises in the early phase of the outbreak when case numbers are low and stochastic effects are largest; for the 2009 pandemic in the UK, this stochastic variability is further amplified by the fact that school summer holidays interrupted transmission in late July. However, we do not believe this issue risks substantially biasing our parameter estimates, as the analyzed epidemiological data come from the periods where transmission was behaving almost deterministically (due to the very large numbers of infected people). Our parameter estimates regarding transmissibility and immunity are driven by the growth rate of incidence seen in the different waves and the overall (age-dependent) attack rates seen, and (conditional on the observed timing of the waves) these are not subject to substantial stochastic variation.

Where we are underestimating the true uncertainty in parameter estimates is in relation to seeding of the first and third waves of transmission. While our central estimate of the number of cases seeding infection in wave 1 would be expected (from Figure S35) to be unbiased, we are almost certainly underestimating the width of its credible interval by not accounting for the highly stochastic nature of transmission when numbers of infectives are small. Since we restart transmission when modelling the third wave, the same issue applies to the estimates of the numbers of infectives introduced to seed the third wave. However, this does not affect our estimates of the growth rate (and thus transmissibility) since when exponential growth is seen, transmission dynamics are (by definition) close to deterministic.

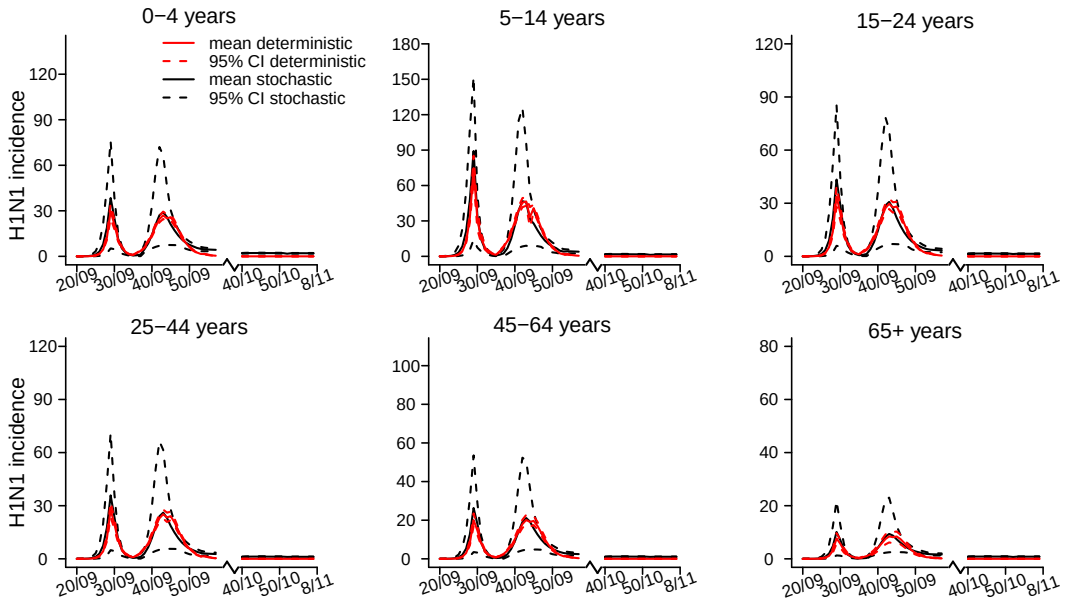


Figure S35: Comparison between deterministic and stochastic realizations of the model obtained by drawing 500 parameters sets from the posterior distribution of model variant 1. For each parameter set we simulate 100 realizations of the stochastic model; the average has been computed on all simulations.

4 Simulating alternative pandemic vaccination policies

In total, the UK government received delivery of 35 million doses of pandemic vaccine [2]. Despite this, on the basis of the epidemiological evidence available at the time [1], vaccination was restricted to clinical risk groups and the under 5s. We explore here the effects of the implementation of larger scale vaccination campaigns (that would have made fuller use of the purchased stockpile) on the emergence of the third wave. We examine 4 scenarios:

- Scenario 1: 30% coverage of the whole population
- Scenario 2: 30% coverage in $<$ 5s and 40% coverage in $>$ 5s
- Scenario 3: 30% coverage in $<$ 5s and 50% coverage in $>$ 5s
- Scenario 4: 30% coverage in $<$ 5s and 60% coverage in $>$ 5s

The coverage levels considered in the four scenarios are consistent with the stockpile of 35 million doses available at the time, assuming that a single dose is sufficient to induce protection [13, 14, 15].
We model vaccination of children under 5 years old as occurring in early April 2010 (see main text) and vaccination of individuals over 5 years as occurring at the start of June 2010. We assume 100% vaccine efficacy in children under 5 years [13], 95% vaccine efficacy for individuals 5-64 years old [14] and 80% vaccine efficacy in those 65 or over [15]. Simulations are obtained by drawing 500 parameter sets from the posterior distribution of model 6 with vaccination fitted on the 2009-2011 surveillance data.
Figures S36-S39 show the H1N1-attributable ILI incidence and seroprevalence predicted by scenarios 1-4 respectively. Figure S40 shows the effective reproduction number and H1N1 incidence over time together with wave-specific attack rates for scenario 1. Figure S41 shows the predicted effective reproduction number over time for scenario 2, 3 and 4.

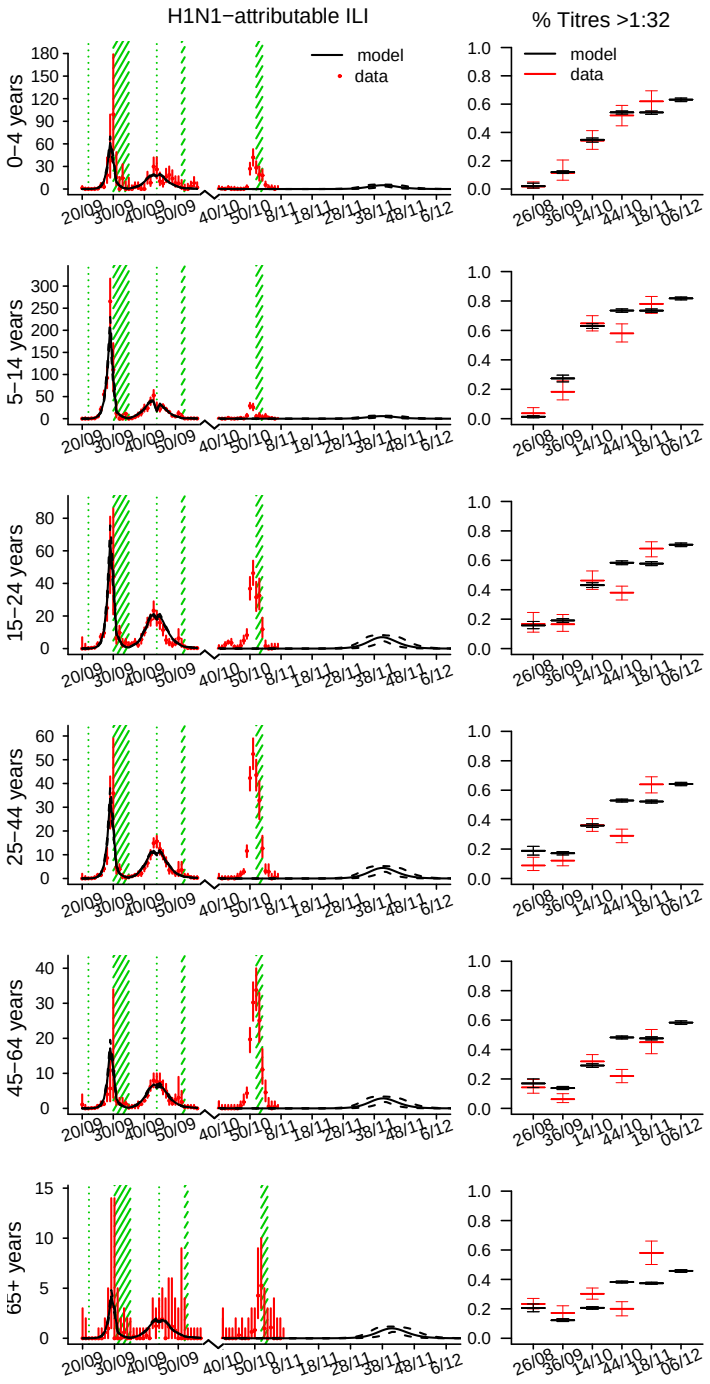


Figure S36: Simulation of model variant 6 assuming 30% vaccine uptake (scenario 1). Left panels: observed (red) and simulated (black) H1N1-attributable ILI incidence (x 100 000) by week. Solid black lines represent the mean, dashed black lines represent the 95% CrI of the simulations. Red bars show the exact 95% CI around the data. Green shaded areas represent holiday weeks. Right panels: observed mean and exact 95% CI (red) and simulated mean and 95% CrI (black) of the proportion of samples showing HI titres $> 1:32$ at the pre-pandemic baseline (week 26/08), after the first wave (week 36/09) and after the second wave (week 14/10).

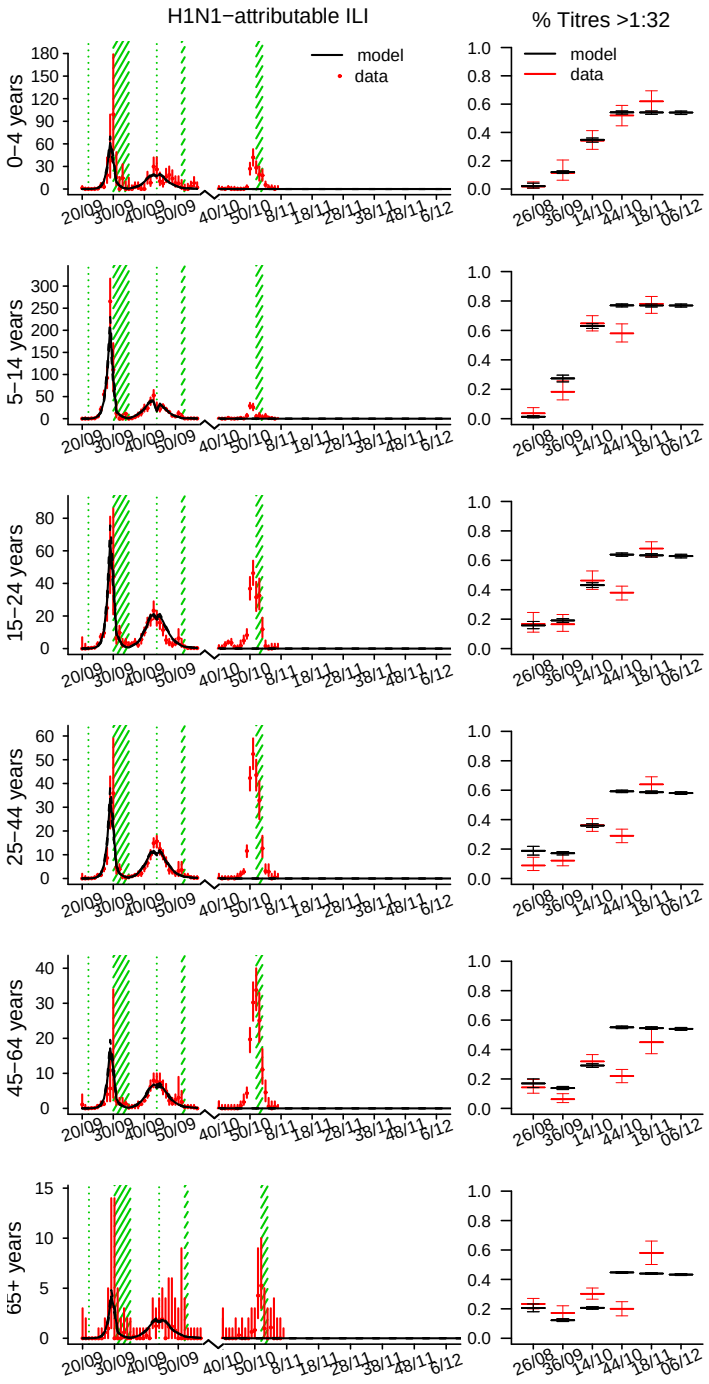


Figure S37: As in Figure S36, but assuming 30% vaccine uptake in under 5s and 40% vaccine uptake in the rest of the population (scenario 2).

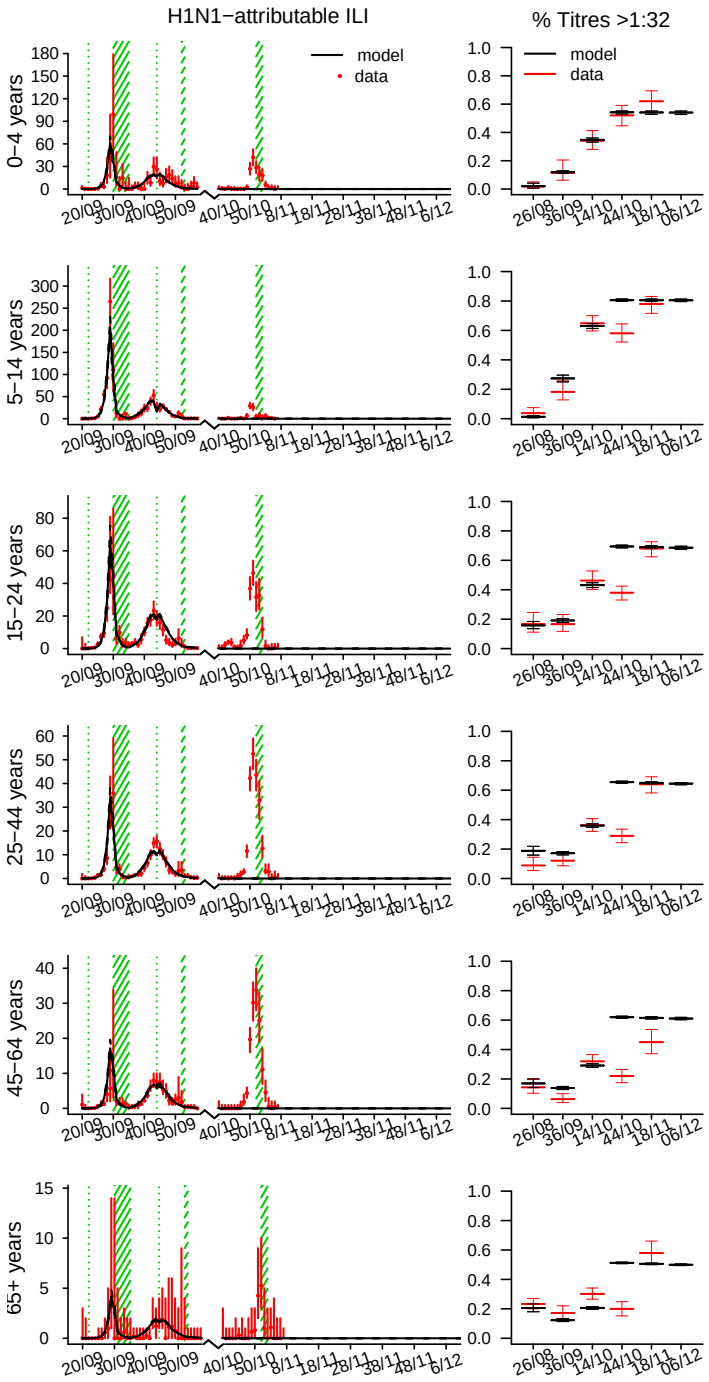


Figure S38: As in Figure S36, but assuming 30% vaccine uptake in under 5s and 50% vaccine uptake in the rest of the population (scenario 3).

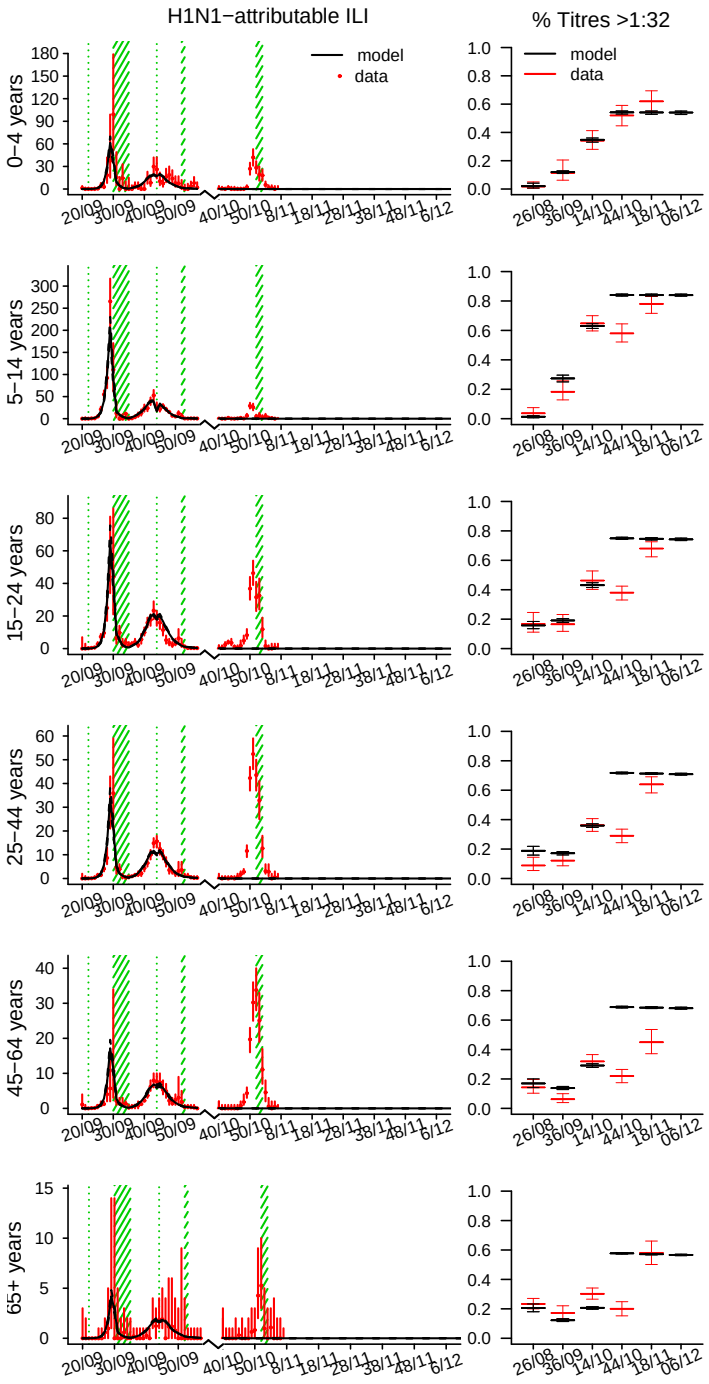


Figure S39: As in Figure S36, but assuming 30% vaccine uptake in under 5s and 60% vaccine uptake in the rest of the population (scenario 4).

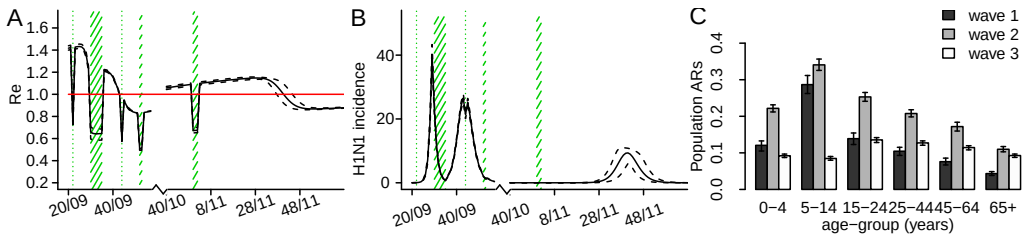


Figure S40: Simulation results for model variant 6 assuming 30% vaccine uptake (scenario 1). (A) Effective reproduction number in time (solid line=mean, dashed lines=95% CrI, green shaded areas=holiday weeks). (B) Incidence of H1N1 cases in the population (x 1000) in time (solid line=mean, dashed lines=95% CrI, green shaded areas=holiday weeks). (C) Attack rates by age group and wave (mean and 95% CrI).

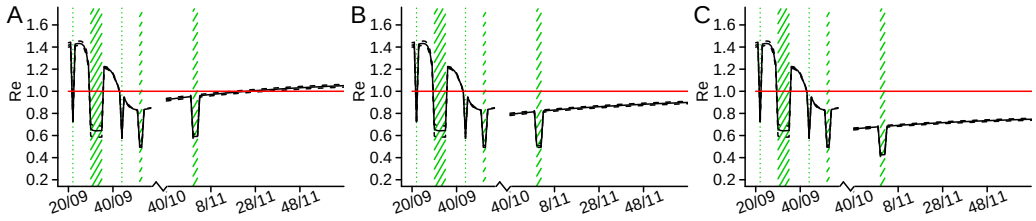


Figure S41: Estimated effective reproduction number in time (solid line=mean, dashed lines=95% CrI, green shaded areas=holiday weeks) obtained with model variant 6 assuming 30% vaccine uptake in under 5s and (A) 40% vaccine uptake in the rest of the population (scenario 2); (B) 50% vaccine uptake in the rest of the population (scenario 3); (C) 60% vaccine uptake in the rest of the population (scenario 4).

5 Climatic factors

There is evidence that climatic factors facilitate virus transmissibility and affect the population susceptibility [16]. Historical climate data for the UK are available on the Met Office website (`www.metoffice.gov.uk./climate/uk/datasets`). In Figure S42 we plot the minimum, mean and maximum temperatures (in degree Celsius) and rainfall (in mm) observed in England in the months of October, November, December and January from 2000 to 2010. Figure S42 shows that December 2010 was one of the coldest and driest Decembers in that decade.

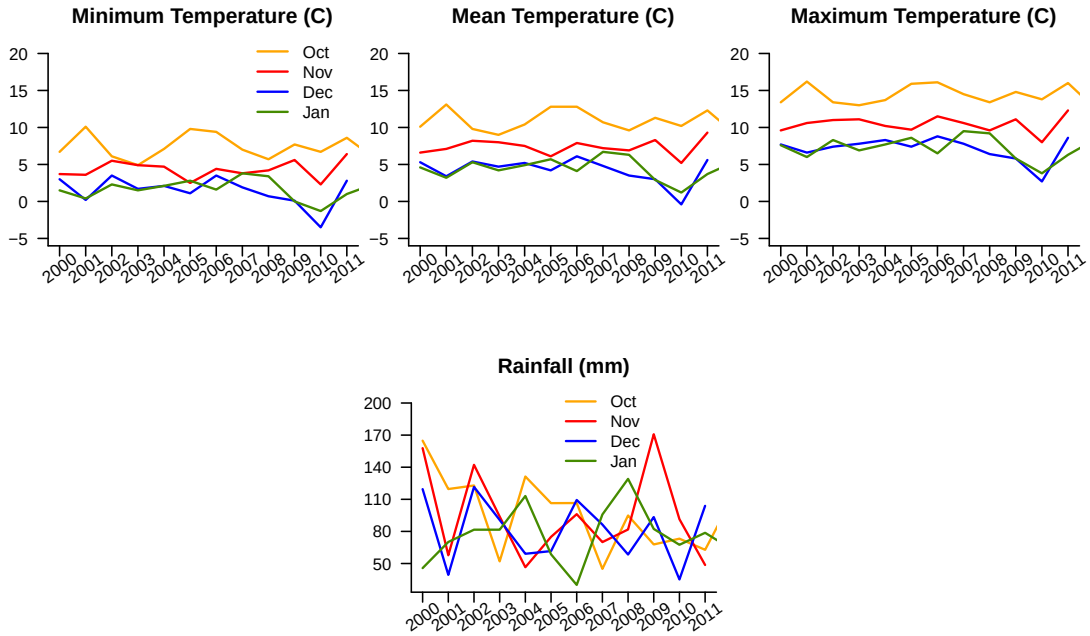


Figure S42: Plot of the historical monthly average temperatures (minimum, mean and maximum in degree Celsius) and rainfall (in mm) recorded in England since 2000 by the Met Office [data accessed on November 15, 2012].

References

[1] Hardelid P et l. (2010) Assessment of baseline age-specific antibody prevalence and incidence of infection to novel influenza A/H1N1 2009. *Health Technol Assess*, 14(55): 115–192.

[2] Hine DD. The 2009 Influenza Pandemic. An independent review of the UK response to the 2009 influenza pandemic. URL `http://webarchive.nationalarchives.gov.uk/+/http://www.cabinetoffice.gov.uk/media/416533/the2009influenzapandemic-review.pdf` [Accessed February 7, 2013]

[3] Department of Health. The H1N1 swine flu vaccination programme 2009-2010. URL `http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_106299.pdf` [Accessed January 28, 2012]

[4] Health Protection Agency. Seasonal influenza vaccine uptake amongst GP patient groups in England. Winter season 2010-11. URL `http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_129856.pdf` [Accessed January 28, 2012]

[5] Health Protection Agency. Pandemic H1N1 (Swine) Influenza Vaccine Uptake amongst Patients Groups in Primary Care in England. URL `http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/@ps/documents/digitalasset/dh_121014.pdf` [Accessed January 28, 2012]

[6] O.Diekmann O and Heesterbeek JAP (2000) *Mathematical Epidemiology of Infectious Diseases*. (Wiley, Chichester, New York)

[7] Press WH, Teukolsky SA, Vetterling WT, and Flannery BP (2002) *Numerical Recipes in C++, The Art of Scientific Computing*. (Cambridge University Press)

[8] Dorigatti I, Cauchemez S, Pugliese A, and Ferguson NM (2012) A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: Application to the Italian 2009-2010 A/H1N1 influenza pandemic. *Epidemics*, 4(1):9 – 21.

[9] Kuske R, Gordillo LF and Greenwood P (2007) Sustained oscillations via coherence resonance in SIR. J Theor Biol, 245: 459-469.

[10] Dushoff J, Plotkin JB, Levin SA and Earn DJD (2004) Dynamical resonance can account for seasonality of influenza epidemics Proc Natl Acad Sci USA, 101(48): 16915-16916.

[11] Black A and McKane AJ (2010), Stocahstic amplification in an epidemic model with seasonal forcing, J Theor Biol, 267: 85-94.

[12] Alonso D, McKane A and Pascual M (2007) Stochastic amplification in epidemics J R Soc Interface, 4: 575-582.

[13] Carmona A et al. (2010) Immunogenicity and safety of AS03-adjuvanted 2009 influenza A H1N1 vaccine in children 6–35 months *Vaccine*, 285837–5844.

[14] Greenberg ME et al. (2010) Response to a monovalent 2009 Influenza A (H1N1) Vaccine *N Engl J Med*, 361(25):-2413.

[15] Zhu FC et al. (2010) A novel Influenza A (H1N1) Vaccine in Various Age Groups *N Engl J Med*, 361(25):-2422.

[16] Shaman J and Kohn M (2009) Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci USA*, 106(9):3243–3248.