Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London

Paul J. Birrell[a,1], Georgios Ketsetzis[b], Nigel J. Gay[c], Ben S. Cooper[d], Anne M. Presanis[a], Ross J. Harris[b], André Charlett[b], Xu-Sheng Zhang[b], Peter J. White[b,e], Richard G. Pebody[b], and Daniela De Angelis[a,b]

[a]Medical Research Council Biostatistics Unit, University Forvie Site, Robinson Way, Cambridge CB2 0SR, United Kingdom; [b]Health Protection Services, Health Protection Agency, 61 Colindale Avenue, London NW9 5HT, United Kingdom; [c]Fu Consulting, Hungerford, United Kingdom; [d]Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok 10400, Thailand; and [e]Medical Research Council Centre for Outbreak Analysis and Modelling, Department for Infectious Disease Epidemiology, Imperial College Faculty of Medicine, Norfolk Place, London W2 1PG, United Kingdom

Edited by Aravinda Chakravarti, Johns Hopkins University School of Medicine, Baltimore, MD, and accepted by the Editorial Board September 26, 2011 (received for review February 28, 2011)

The tracking and projection of emerging epidemics is hindered by the disconnect between apparent epidemic dynamics, discernible from noisy and incomplete surveillance data, and the underlying, imperfectly observed, system. Behavior changes compound this, altering both true dynamics and reporting patterns, particularly for diseases with nonspecific symptoms, such as influenza. We disentangle these effects to unravel the hidden dynamics of the 2009 influenza A/H1N1pdm pandemic in London, where surveillance suggests an unusual dominant peak in the summer. We embed an age-structured model into a Bayesian synthesis of multiple evidence sources to reveal substantial changes in contact patterns and health-seeking behavior throughout the epidemic, uncovering two similar infection waves, despite large differences in the reported levels of disease. We show how this approach, which allows for real-time learning about model parameters as the epidemic progresses, is also able to provide a sequence of nested projections that are capable of accurately reflecting the epidemic evolution.

Bayesian statistics | real-time modeling | general practice consultation data | infectious disease | seroepidemiology

An emerging epidemic engenders an increased demand upon health services. Resolving the extent to which this is due to high levels of disease transmission as opposed to a heightened public sensitivity is essential for determining the appropriate public health response.

This was especially crucial when estimating the course of the 2009 influenza A/H1N1pdm outbreak in England, where, unusually, the pandemic resulted in a summer peak in rates of consultation at general practices (GPs) for influenza-like illness (ILI). This is clearly demonstrated by data from the return service of the Royal College of General Practitioners (RCGP) in Fig. 1*A* where weekly GP consultation rates per 100,000 population over the 2009 pandemic are compared with rates from the three previous years. Also shown is the proportion of swabbed individuals whose swabs tested positive for the presence of any flu virus (*SI Data*). Note that the GP consultation rate for 2009 is much higher than the usual seasonal rate, whereas the corresponding positivity is comparable to that observed in the preceding winters. This suggests that a substantial proportion of the peak in consultations was not directly attributable to A/H1N1pdm. Conversely, serological studies (3) have shown a marked increase in the prevalence of influenza antibodies among the population. Therefore, the degree to which the increased demand upon GPs is due to high levels of disease transmission as opposed to heightened public sensitivity remains unclear (4). Fig. 1 *B* and *C* show GP consultation rates by region and age group: consultations in Greater London and the West Midlands exhibit rapid early exponential growth, but the peak in London is much higher; rates appear to decrease markedly with age. Importantly, a first peak occurs immediately prior to the summer school holiday and the launch of the National Pandemic Flu Service (NPFS) phone line, introduced to relieve the pressure on GPs and expedite antiviral distribution (see *SI Data*); a second, much smaller peak, is observed in the autumn. This evidence, supported by the work of ref. 5, promotes the further hypothesis of a fluctuating propensity for individuals with symptoms of ILI to seek medical attention, perhaps induced by media coverage and changing governmental advice, as well as the social distancing effects of school holidays.

Traditionally, transmission modeling is used to investigate epidemic development. In the area of infectious respiratory diseases, many approaches have been proposed (6–13), including those that account for the effects of behavioral changes upon transmission (14), explicitly model the impact of school closure (15), and incorporate a temporally varying case-detection rate (16).

Here we model these aspects simultaneously while additionally accounting for the time-varying noise in the data due to consultations for non-A/H1N1pdm ILI. This is achieved by developing a model for integrating noisy GP consultation data, virological positivity data, virologically confirmed case data, and information from serological (seroprevalence) surveys (see *SI Data*). Each dataset is available and used at daily intervals. An age-structured transmission model is embedded within a Bayesian framework, allowing incorporation of any a priori information about model parameters from previous influenza strains via probability distributions. These prior distributions are then updated by available data to provide posterior statements about parameters of interest and their uncertainty, presented here in the form of 95% credible intervals (CrIs).

Results

Fig. 2 is a schematic representation of the model used to describe the data-generating process. Three different components are knitted together: an age-structured transmission-governing component, a disease component, and a third component describing the mechanisms through which infected individuals report their symptoms to the health-care system. In the transmission component, susceptible individuals (S) become exposed (E) through an effective contact with infectious (I) individuals and become infective themselves after a short latent period, to be then removed (R) from the pool of infectious individuals after a further period. Transmission is governed both by a time- and age-varying force of infection $\lambda(t,a)$, depending on the transmissibility of

Author contributions: N.J.G., B.S.C., and D.D.A. designed research; P.J.B., G.K., and X.-S.Z. performed research; P.J.B., R.J.H., A.C., P.J.W., and R.G.P. analyzed data; and P.J.B., A.M.P., and D.D.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.C. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: paul.birrell@mrc-bsu.cam.ac.uk or daniela.deangelis@mrc-bsu.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103002108/-/DCSupplemental.

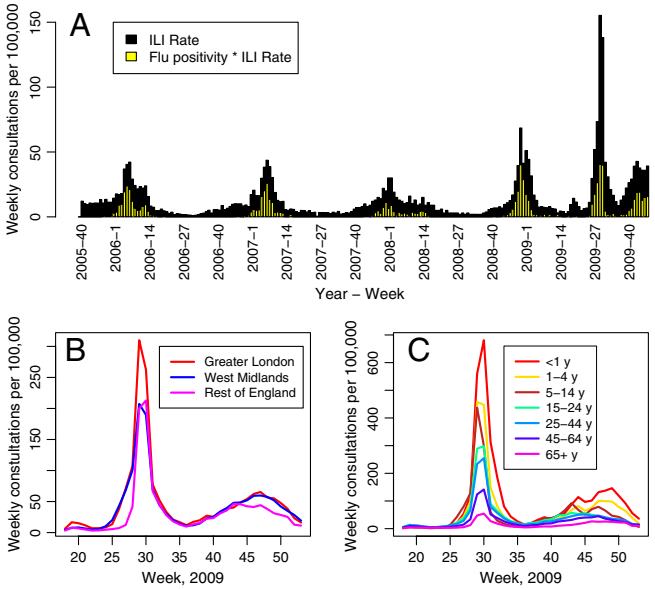

Fig. 1. (*A*) Time series of GP consultation rate for ILI and virological positivity within the RCGP surveillance scheme in England and Wales (1), by week, from end-2005 to end-2009. The positivity is given by the proportion of each bar shaded yellow. (*B*) Observed weekly consultation rates per 100,000 within each group (see ref. 2 and *SI Data*, section 1.1) by region (Greater London, West Midlands, and Rest of England) and (*C*) by age group, over the weeks 18–53, 2009.

the virus and the mixing patterns in the population and by the transition rates among the S, E, I, and R states (see *Material and Methods*). Only a proportion, θ, of the newly exposed individuals develop febrile symptoms, from which further proportions $p_{GP}(t,a)$ and p_{CC} consult their GP or have their illness virologically confirmed. Note that $p_{GP}(t,a)$ is calendar time and age-specific to accommodate potential fluctuations in consultation behavior. There are no direct data on transmission. However, serological surveys (see *SI Data*), carried out before and during the epidemic, provide data on indicators, $Z(t,a)$ (see *Material and Methods*), informing the level of susceptibility within the population at the epidemic onset and over time. We make direct observation of $X_{CC}(t,a)$, the symptomatic cases that are virologically confirmed, though this is limited to the early stages of the epidemic (see *SI Data*). Only *indirect* information is available on the number of symptomatic cases consulting GPs, $X_{GP}(t,a)$, in the form of routine surveillance (see *SI Data*, section 1.1) counts of GP consultations for *all* ILI. This includes a background

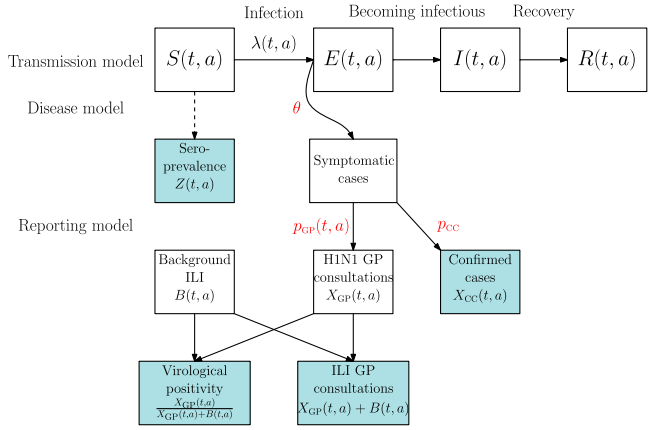

Fig. 2. Model schematic diagram representing the data-generating process. The shaded boxes represent the quantities upon which we make observation.

component, $B(t,a)$, of non-A/H1N1pdm ILI. To identify these two components we use the total number of ILI consultations $Y(t,a) = X_{GP}(t,a) + B(t,a)$ and information on the virological positivity $X_{GP}(t,a)/(X_{GP}(t,a) + B(t,a))$ (see *Material and Methods*). By combining direct, indirect, and prior information, we produce posterior distributions for the process governing parameters (see *Material and Methods*) and other quantities of interest.

Reconstructing the Epidemic. Fig. 3*A* shows the posterior median and pointwise 95% CrI for the total number of weekly incident infections of A/H1N1pdm in Greater London, using 245 d of epidemic data covering May 1 to December 31 (i.e., from week 18 to week 53) of 2009. Additionally, Fig. 3*A* also shows the estimated age-specific incidences. Much like the GP consultation data, the epidemic occurs in two waves: a summer first wave (May to end-August) and an autumn second wave (September to December). The first wave rises sharply to a peak of 109,000 (81,000–146,000) new infections in the week immediately prior to the school holidays. The second wave has a smaller peak with posterior probability 0.885. Conversely, as can be seen from Table 1, which reports estimates of the infection attack rate (i.e., the cumulative incidence expressed as a proportion of the total population), there is slightly larger cumulative incidence in the second wave, a phenomenon not at all evident from the GP consultation data (Fig. 1 *B* and *C*). The discrepancy between the GP consultation data and the estimated infection pattern is clarified in Fig. 3*B*, which compares the cumulative consultations with the estimated cumulative infections, both calculated as proportions of their corresponding total. In this plot, a steep gradient identifies points in time during the pandemic when a relatively high density of GP consultations (or infections) occur. From the steep gradient in the GP consultations curve over weeks 27 and 28, it can be seen that the consultations were highly localized around this time. The separation between the two lines and the smaller

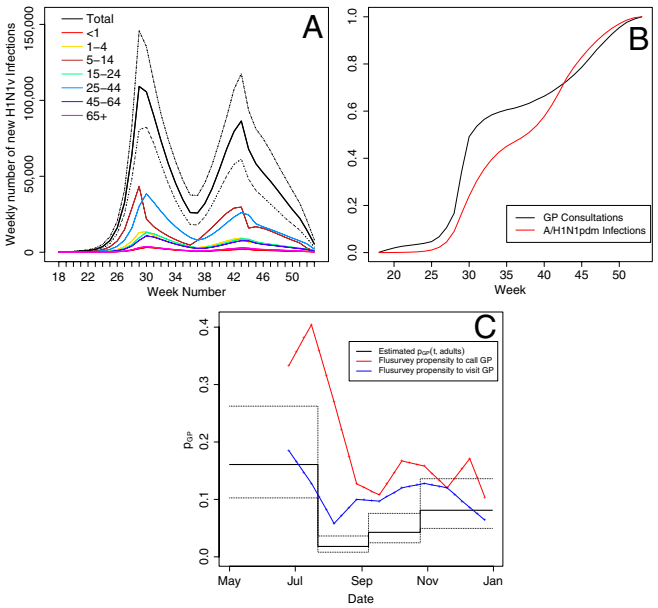

Fig. 3. (*A*) Estimated weekly infections for Greater London, spanning weeks 18 to 52 of 2009, as reconstructed from the model. The black line is the total incidence of infections over all ages and the dotted lines represent a 95% CrI. (*B*) The cumulative incidence of GP consultations (black line) versus the estimated (posterior median) cumulative incidence of infection (red line), both expressed as a fraction of the respective total incidence. (*C*) The estimated propensity for incident adult symptomatic cases to consult with a GP over time (black line). The red and blue lines present the propensity for ILI cases to call and visit (respectively) their GP from an Internet-based survey, FluSurvey (*SI Materials and Methods*).

Table 1. Posterior median and 95% CrI for cumulative incidence of infections and cases and age-specific attack rates

	End-August	September–December
Infections	672,362 (544,466, 826,014)	784,934 (612,369, 981,410)
Cases	217,003 (133,434, 324,027)	257,105 (160,860, 385,760)
Attack rates, %		
(Overall)	8.9 (7.2, 10.9)	10.4 (8.1, 13.0)
<1 y	15.3 (12.4, 18.9)	14.9 (12.1, 18.2)
1–4 y	19.5 (15.1, 26.1)	20.8 (16.9, 25.4)
5–14 y	22.1 (18.0, 26.7)	29.8 (24.1, 35.3)
15–24 y	7.9 (5.9, 10.5)	8.6 (6.3, 11.4)
25–44 y	8.0 (6.2, 10.4)	9.0 (6.7, 11.7)
45–64 y	4.1 (3.1, 5.5)	4.7 (3.4, 6.3)
65+ y	2.5 (1.9, 3.3)	3.0 (2.2, 4.0)

gradient of the infection curve showed that, unlike the GP consultations, infection is shared out more evenly over the two waves.

Returning to Table 1, children, i.e., individuals younger than 15 y old, acquired the most disease: approximately 52% of 5–14, 40% of 1–4, and 30% of under 1 y-old children are estimated to have contracted the virus, substantially higher than the overall infection attack rate of 19%. Note that Fig. 1 shows that GP consultation rates decline with age. In contrast, the estimated attack rates peak in the 5–14 age group, indicating a greater component of background consultation in the <5 s. The precipitous decline in the infections brought about by the school holidays (both peaks of Fig. 3*A* occur in the same week as the start of a school holiday) highlights the key role that children play as agents of transmission, as seen in estimates of scaling factors that modify contact rates (parameters m_i in Table 2). Compared to school term time, we estimate a reduction in the rate of contact within the 5–14 age group of 72%(52%–97%) in the summer holiday ($1-m_3$) and of 48%(22%–72%) in the half-term school holidays ($1-m_5$). See *Materials and Methods* for further details. The data are, however, unable to identify a similar effect among the 1–4 y-olds (see the wide CrI attached to parameters m_2 and m_4). We further estimate that child-to-child infectious contacts are 2.13(1.86–2.47) ($=1/m_1$, Table 2) times as likely to result in transmission than those involving at least one adult. The effect of this estimated fall in contact rates in the summer holiday and the contribution of children to transmission translates into a reduction of 35.2% (30.2%–40.2%) in the effective reproductive number: the average number of secondary infections induced by a primary infection at a given time. In a fully susceptible population this reduction would be similar: 36.4%(30.9%–41.6%).

Also from Table 2 we can see that the proportion of infections that develop into symptomatic cases, θ, which has an informative prior (see Fig. 4*B*), is estimated to be 0.33(0.21–0.47). This corresponds to around 35,000 incident symptomatic cases at the peak of the first wave. The posterior median for the basic reproductive number, R_0, is 1.65(1.56–1.75). As with θ, R_0 shows considerable prior to posterior divergence, whereas the posterior for the mean infectious period, d_I, is nearly identical to the prior (see Fig. 4*B*).

At NPFS launch, the propensity for adults to consult is estimated to fall from 16% to 1.8% (Fig. 3*C*). Only a small increase follows at a second breakpoint in early September, but by a third

Table 2. Posterior median and 95% CrI for key parameters

Parameter	Estimate	(95% CrI)
R_0	1.65	(1.56, 1.75)
d_I	3.45	(3.32, 3.57)
m_1	0.469	(0.405, 0.539)
m_2	0.535	(0.0395, 0.970)
m_3	0.279	(0.0323, 0.481)
m_4	0.295	(0.0115, 0.919)
m_5	0.522	(0.276, 0.784)
θ	0.328	(0.211, 0.468)

Key parameters are R_0, the basic reproductive number; d_I, the average infectious period (in days); θ, the proportion symptomatic; m_i, scaling factors applied to the mixing matrices to account for relative infectivity (m_1) and school holiday effects

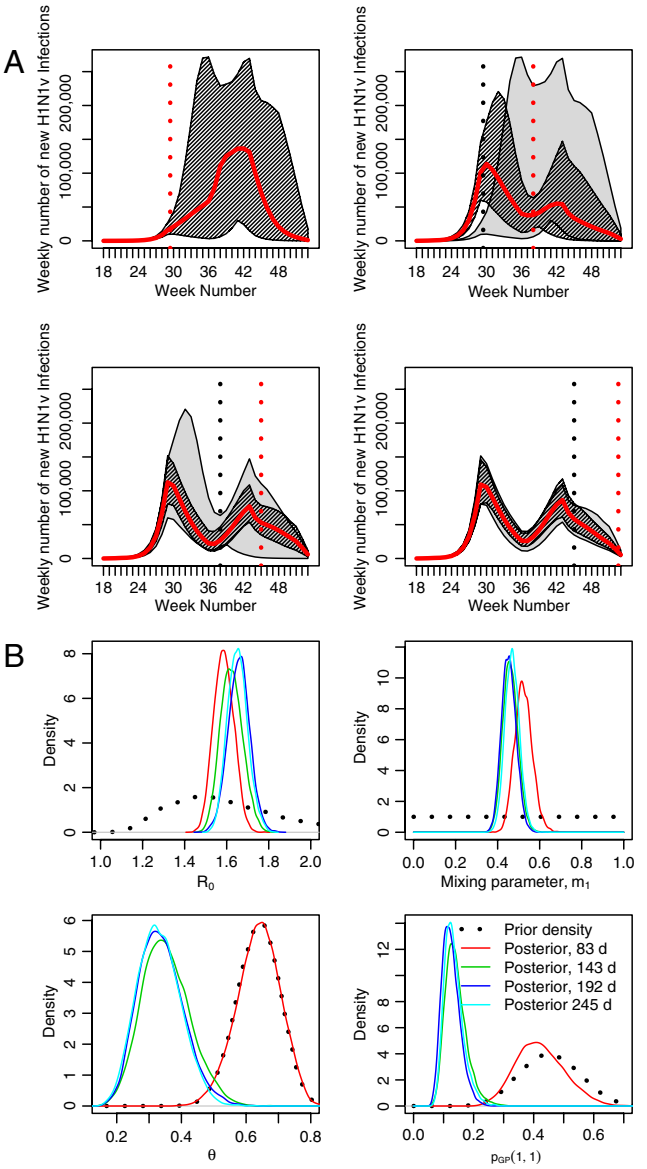

Fig. 4. (*A*) Sequential epidemic reconstructions/projections based on 83, 143, 192, and 245 d of surveillance data. The gray shaded area shows the 95% CrI for the epidemic construction from the temporally previous analysis, with the darker shaded area showing the "current" analysis. The solid red line indicates the posterior median number of infections, the dotted red vertical line shows the time at which the surveillance data ends, and the dotted gray vertical line shows the time at which the previous batch of surveillance data ended. (*B*) Sequentially obtained posterior density estimates, plotted alongside corresponding prior distributions for the parameters R_0, m_1, θ, and $p_{GP}(1,1)$.

breakpoint, in late October, the propensity returns to a value close to 10%. Similar results are obtained for this parameter in children. These estimates are similar to values expected during seasonal influenza epidemics (17, 18), but lower than estimates from the Internet-based Flusurvey (see Fig. 3*C* and *SI Data*), possibly reflecting biases in the population captured by the survey.

Predicting the Epidemic. The above results are related to an epidemic that is now over. A crucial question is whether the model can be used as a tool for inferences and predictions while an epidemic is ongoing. To assess this, further analyses were conducted based on 83, 143, and 192 d of epidemic surveillance data. The 83-d analysis contains no serological data except those used to inform the baseline prevalence of antibodies, whereas the 143- and 192-d analyses incorporate serological data collected during the epidemic (see ref. 3 and *SI Data*).

Fig. 4*A* illustrates how the predictions evolve as data accumulate. Fig. 4*B* shows how the estimated posterior densities for the parameters R_0, m_1, θ, and $p_{\mathrm{GP}}(1,1)$, the consultation propensity in children in the first 83 d (see *SI Materials and Methods*), evolve over time, starting with their prior distributions. From the 143-d analysis onward, credible intervals for the future number of infections appear to enclose the estimated numbers in the subsequent analysis. However, this is not so in moving from the 83- to the 143-d analysis. This is due to the lack of serological data in the 83-d analysis. Given the large degree of dispersion, the GP consultation data are too weakly informative to overcome the informative priors placed, partly for the sake of identifiability, upon parameters such as θ and $p_{\mathrm{GP}}(\cdot\,,\cdot)$ (see *SI Materials and Methods*). The densities of Fig. 4*B* show that, in the earliest analysis with no serological data, the posterior distributions for these parameters are near identical to the priors, centered on values far larger than the posteriors obtained from the subsequent analyses. The inclusion of the serological data in the 143-d analysis provides a clear indication of the level of cumulative incidence, which is higher than the 83-d results might suggest. With stronger information on the incidence, the data become sufficiently informative to overcome the prior distributions for θ and p_{GP}, and this is reflected in a shifting of the posterior distributions to values that remain consistent across the 143-, 192-, and 245-d analyses.

Discussion

Our approach allows reconstruction and projection of the trajectory of an epidemic by disentangling epidemic and behavioral dynamics. Combining data from different sources is crucial, as each plays an important role: virological data partition consultations between A/H1N1pdm and other ILI, GP consultations determine the temporal trend, and the serological data give the scale of the epidemic.

Our estimates of attack rates are lower than has been obtained elsewhere (19), importantly providing improved understanding of how a third wave of A/H1N1pdm infection occurred in late 2010. Although our estimate of R_0 is consistent with that obtained by others (13, 16), the estimate that only a third of infections are symptomatic is much lower than corresponding estimates from Mexico and Hong Kong [0.86 and 0.64, (13, 20)] but comparable to estimates from New Zealand [0.45 (21)] and France [0.20 (22)] and in broad agreement with the systematic review of ref. 23. Our estimate of $0.47(=1/2.13)$ for the relative risk of transmission in infectious contacts involving at least one adult is in direct agreement with a previous estimate of 0.485 (95% CrI 0.302–0.625) (13). By combining our estimated θ and attack rates, we obtain a number of symptomatic cases, which is a fourfold increase on the central and a twofold increase on the upper bound of the official estimates for the two waves (24). Previous work (10) uses these central estimates as data, multiplying them by a factor of 10 in order to achieve a good model fit. This factor can be interpreted as a product of two components: one that accounts for the asymptomatic infections ($1/\theta$) and one that accounts for underascertainment in the symptomatic case number estimates. Here, these two components multiply to give a factor of approximately 12.

In its transmission component, the model is similar to that used elsewhere (16), where data on laboratory-confirmed cases, modeled using a time-varying reporting rate, have been considered. Here, in addition to some lab-confirmed cases, other sources of information are used, notably noisy GP consultation data. This enables us to advance earlier efforts to model the epidemic in England. Using a hybrid of estimation approaches, ref. 10 treats estimated weekly incident cases of A/H1N1pdm (24) as data, with no propagation of the error inherent in the estimation process. We employ a more rigorous, statistical approach, which utilizes a richer array of raw data. For the same pandemic in Singapore, ref. 25 implements an algorithm for online updating of estimates arising from an S, E, I, R transmission model fitted to GP consultation data alone. This approach, which makes no stratification by age, suffers similarly from a lack of nesting in the early stages and masks the interage group transmission dynamics. In the same vein, ref. 18 develops a methodology for real-time inference, but this offers little opportunity to make any learning about many key model parameters. Our model also allows for the quantification of the impact of school holidays, the level of non-A/H1N1pdm consultation, and obtains estimates for the propensity for patients affected by the A/H1N1pdm virus to seek consultation.

Assumptions. One key strength of any model used to predict or assess epidemic impact is a robustness to (often unavoidable) modeling assumptions. In *SI Sensitivity Analyses* we investigate the impact of dropping, or changing, a number of the assumptions made in producing the epidemic estimates obtained here. Our results are generally robust to reasonable deviations from modeling assumptions, yet also suggest avenues for further investigation. Specifically, we examine three modeling components: the performance of the virological testing procedure, the assumed contact patterns, and the assumed functional form of the propensity to consult, $p_{\mathrm{GP}}(t,a)$.

Test Sensitivity. Thus far, it has been assumed that the virological testing procedure has a sensitivity of 1; i.e., there are no false negatives. If we relax this assumption to reasonable values for the test sensitivity, say 0.8 and 0.9, the key results change very little. As expected, there is a small increase in the estimated symptomatic attack rates, as a lower sensitivity allows for more A/H1N1pdm cases among the GP consultations, but there is negligible impact on the total infection attack rates. Of the values investigated, a test sensitivity of 0.8 had larger values for the likelihood in its posterior distribution.

Mixing Matrices. The mixing matrices used to describe rates of contact between the different age groups are based upon United Kingdom (UK) data from the POLYMOD (Improving Public Health Policy in Europe through Modeling and Economic Evaluation of Interventions for the Control of Infectious Diseases) study (see ref. 26). Results are robust to small changes in both parameterization of the mixing matrices (see *Materials and Methods*) and the rates of contact themselves. More interestingly, a preliminary attempt to estimate the entire mixing matrices (*SI Sensitivity Analyses*), using the POLYMOD data only as prior information, indicates that contacts among the 5–14 age group are particularly important, while also suggesting that contacts between this age group and the 15–24 age group may be more influential than previously thought. Adopting these estimates results in a slight fall in the estimated R_0, while increasing the total infection attack rate. This is due to a shift in the infection profile from small children to the bigger pool of susceptibles in the 15–24 age group. Developing a more rigorous approach to

estimation of contact patterns, and in particular the choices of the informative priors, constitutes a promising avenue for future work.

Propensity to Consult. A piecewise linear parameterization for $p_{GP}(t,a)$ in the post-NPFS era gave results not materially different from those featured here. Attempts to adopt this piecewise linear parameterization over the entire epidemic period resulted in lack of identifiability and undue influence of prior distributions (see *SI Sensitivity Analyses*).

Further Applicability. Our modeling approach has focused on the reconstruction of the epidemic in a globally prominent metropolitan region. However, the general methodology presented here is highly applicable to any influenza epidemic within any similar health-care network. Although it is unreasonable to carry out this modeling exercise on England as a whole, due to the unrepresentativeness of the pooled virological and serological data, we have repeated the analyses in three other disjoint regions, which, together with London cover the whole of England (see *SI Further Applicability*). London and West Midlands experience very similar epidemics, whereas the remainder of the country is split into two regions, North and South, neither of which had a substantial first wave of infection. There is mostly nonsignificant variation in the estimated parameters across the four regions, with the exception of R_0, which suggests a possible link between the reproductive number of the epidemic and the population density within the affected region (London and West Midlands are England's two most densely populated areas).

Serological Data. As epidemic surveillance data accumulate over time, the model is capable of producing sequential epidemic estimates that converge (in the sense that successive credible intervals are nested) and could be used for real-time modeling and prediction. However, we have shown surveillance data alone are insufficient unless or until the epidemic is very far progressed. In general, to generate reliable projections early in the epidemic, the timely availability of relevant data on cumulative incidence and/or the proportion of infections reported in surveillance (a role originally envisaged for the Flusurvey during England's 2009 A/H1N1pdm outbreak; see *SI Data*) is required. Serological data, in particular, are shown in this paper to be vital to ensure convergence of sequentially obtained estimates, given our choice of priors. Analysis conducted in the absence of serological data (see *SI Serological Studies*) shows that with surveillance data alone, a realistic epidemic reconstruction is still impossible at 192 d, i.e., after the peak of the second wave. As a result, any online inference is rendered highly infeasible. This highlights the critical importance of the timely availability of serological information in an emergent epidemic, when information on key parameters may be lacking and/or priors may be misspecified, as here. This is clearly a challenge, given current limitations on test developments, facilities, and recruitment of appropriately representative populations (3, 19), but one that is very important to meet.

Materials and Methods

Our approach integrates data from a number of sources, combining information from GP surveillance networks with epidemic specific data. The *SI Data* section provides an in-depth description of the available data on GP consultations for ILI, virological positivity, virological confirmed cases, and serological surveys.

An Integrated Model. The proposed model in Fig. 2 comprises a transmission model and a disease and reporting model. The model dynamics are deterministic and discrete. We model from the period May 1 to December 31, 2009, and for the following age groups: <1 y, 1–4, 5–14, 15–24, 25–44, 45–64, and 65+ y. The epidemic is initiated with a small number of infectious individuals and a pool of susceptible individuals. At each subsequent time point the transmission model generates a number of newly infected individuals, which enter the disease and reporting model, while the pool of susceptible individuals diminishes. The disease and reporting models then govern the proportion of these incident infections that appear in the GP consultation and confirmed case datasets and the delay inherent in doing so.

In the age-structured S, E, I, R model, transmission is dictated by a time- and age-varying force of infection $\lambda(t,a)$ and transition rates σ and γ, which describe the rates of transition between states $E \to I$ and $I \to R$, respectively. These rates are functions of the mean latent period, d_L, and the mean infectious period, d_I, the expected times spent in states E and I, respectively. The force of infection depends on two key quantities: the basic reproduction number of the virus, R_0, and the relative rates of contact between the different age groups, introduced through the time-varying matrix, $\boldsymbol{M}(t)$. Details of how these quantities combine to give the incident number of infections can be found in *SI Materials and Methods*, Eq. 7. A proportion, θ, of the exposed individuals become clinical cases, with further fractions $p_{GP}(t,a)$ and p_{CC} of these symptomatic individuals consulting their GP or being virologically confirmed, respectively. Typically, there will also be a time lag from infection to either of these events. This delay is assumed to be distributed as a gamma random variable and arises from three independent processes: the incubation time until symptom onset, the delay in reporting the GP consultation or having illness virologically ascertained, and the subsequent reporting delay. These component delays are assumed to have known mean and variances, which are summed to give the mean and variance of the distribution governing the overall time from infection to each event. This is discussed in more detail in *SI Materials and Methods*. The size of the initially susceptible population within an age group, $S(0,a)$, is informed by baseline serological data from 2008 (3). Subsequently, for serological data taken at time t, the expected seropositivity is given by $1 - (S(t,a)/N_a)$, where N_a is the size of the population in age group a.

Modeling Challenges. ***Consultation Behavior.*** The propensity of individuals to consult with their GP given symptomatic ILI varied significantly over the course of the study period. Initially, this propensity was high, as seen from the marked increase in the consultation rates during the first wave, with only a modest increase in the accompanying virological positivity. However, government advice that patients were to consult through the NPFS, rather than their GP, drastically reduced this propensity. The model has to be sufficiently flexible to account for this, as well as permitting some temporal variation in the levels of adherence to the governmental guidelines over time. This impacts upon our model in two ways: (*i*) through the propensity to consult with a GP conditional upon symptomatic infection with A/H1N1pdm, $p_{GP}(t,a)$; this is modeled as a piecewise function over time, with differing rates for children and adults, the details of which can be found in *SI Materials and Methods*; (*ii*) through the "background" consultation of non-A/H1N1pdm patients with ILI symptoms. The background component of the consultation, $B(t,a)$, is parameterized as a piecewise constant function over time, with varying rates for each age group, thus allowing for temporal fluctuation in the behavior and prevalence of individuals with non-A/H1N1pdm ILI. See *SI Materials and Methods* for further details of the model and the estimation of these background rates of consultation, using informative priors derived on the basis of pandemic data heralding from other regions of England.

Mixing Rates and School Holidays. Estimated contact rates based on UK weekday data from within the POLYMOD study (26) formed the basis of the contact matrices $M(t)$ used in our analysis. In school term times, these POLYMOD matrices were modified through the introduction of a scaling factor, m_1, applied to all matrix elements representing a contact rate involving adults. This confers an interpretation upon m_1 of a relative infectivity of adult infectious contacts in comparison to those solely involving children. Effects of school holidays upon disease transmission were accounted for by introducing further factors, m_2 and m_3, which describe the proportionate reduction in rates of contact among 1–4 and 5–14 y-olds, respectively, during the summer holiday. During the shorter half-term holidays, additional multipliers, m_4 and m_5, were applied to the same contact rates to permit differing effects of social distancing brought about by the two types of holiday. See *SI Materials and Methods* for further details.

Inference. ***Parameters.*** Inference is carried out within the Bayesian framework, based upon the posterior distributions of parameters and derived quantities of interest, obtained through the combination of the prior distributions and the likelihood function. We estimate posterior distributions for parameters R_0, d_I, m_i ($i = 1,\ldots,5$), and the size of the initial spark of infection. Conversely, the mean latent period d_L is assumed to be known. Preliminary attempts to estimate both d_L and d_I highlighted that only their sum, not the individual components, is easily identified and these findings have been

formalized elsewhere (27). For the disease and reporting models, we estimate θ, p_{CC}, and the parameters describing $p_{GP}(t,a)$. Furthermore, we estimate the nuisance parameters used to model $B(t,a)$.

Likelihood. If we denote the collection of all model parameters by the vector φ, and

i. w_{ta} is a realization of $W(t,a)$, the virological positivity at time t in age group a, based on a sample of size m^v_{ta}.
ii. x^{CC}_{ta} is a realization of $X_{CC}(t,a)$, the number of lab-confirmed cases at time t in age group a.
iii. y_{ta} is a realization of $Y(t,a)$, the number of GP consultations at time t in age group a.
iv. z_{ta} is a realization of $Z(t,a)$, the seropositivity at time t in age group a, based on a sample of size m^s_{ta}.

Then, treating the above as independent data, the likelihood is given by

$$L(\varphi) = \prod_{t=1}^{n_t}\prod_{a=1}^{n_a} = \{L(w_{ta}|\varphi) \times L(x^{CC}_{ta}|\varphi) \times L(y_{ta}|m^v_{ta},\varphi) \times L(z_{ta}|m^s_{ta},\varphi)\},$$

where n_t and n_a are the number of time points (245) and age groups (7), respectively. The third term in the product gives the likelihood of the GP consultation data. This is modeled through a negative binomial distribution to account for the overdispersion in the count data. This overdispersion is in part due to the within-week pattern of consultation characterized by very few consultations on weekends or bank holidays and a higher rate of reported consultations on Mondays, gradually declining through the week until a small increase on Fridays. The negative binomial distribution is parameterized in terms of the mean number of consultations (as found in *SI Materials and Methods*, Eq. 8) and a piecewise constant dispersion parameter, with one breakpoint at the time of NPFS launch. Otherwise, the confirmed cases, $\boldsymbol{x}_{CC}$, are modeled as Poisson count data, and the positivity and serological data are both treated as realizations of binomial random variables with known denominators.

Priors. A list of the model parameters comprising φ can be found in *SI Materials and Methods*, section 2.3.2. Where possible, parameters have been included as stochastic quantities; i.e., we have placed a prior upon them, so that we can learn about them through the data and so that the modeling procedure incorporates as much a priori knowledge/uncertainty as possible. Some parameters, due to reasons of identifiability, are held to fixed values. Fixed values and the majority of prior distributions are taken from the literature (see *SI Materials and Methods* for details). Such information is deemed to be unknown or unavailable for the parameters of the mixing matrices, m_i, and the overdispersion parameters, and so we place priors that are reasonably uninformative upon them.

Implementation. Posterior distributions for the unknown parameters are evaluated through Markov chain Monte Carlo methods, using a random walk Metropolis algorithm (28, 29). The algorithm was implemented using a bespoke C++ code specifically generated for this class of models. Two separate chains, each consisting of 450,000 iterations, were run in parallel, with the results presented based on a thinned subsample of the final 250,000 iterations from the two chains.

ACKNOWLEDGMENTS. The authors thank the Health Protection Agency Pandemic Influenza team for the timely availability of data; Professor E. Miller for providing serological data; the RCGP Research and Surveillance Centre; the University of Nottingham, Egton Medical Information Systems (EMIS), and EMIS practices contributing to the QSurveillance database. P.J.B., A.M.P., and D.D.A. were funded by the UK Medical Research Council (Grant G0600675). D.D.A. was funded also by the UK Health Protection Agency, as were B.S.C., R.J.H., A.C., X.S.Z., P.J.W., and R.G.P. G.K. was supported by the Commission of the European Community under the Sixth Framework Program Specific Targeted Research Project, SARS (Severe Acute Respiratory Syndrome) Control "Effective and Acceptable Strategies for the Control of SARS and new emerging infections in China and Europe" (Contract SP22-CT-2004-003824). B.S.C. acknowledges support by the Oak Foundation. P.J.W. thanks the Medical Research Council Centre for funding.

1. Fleming DM (1999) Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Public Health* 2:96–100.
2. Harcourt SE, et al. (2011) Use of a large general practice syndromic surveillance system to monitor the progress of the influenza A(H1N1) pandemic 2009 in the UK. *Epidemiol Infect*, 10.1017/S095026881100046X.
3. Miller E, et al. (2010) Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study. *Lancet* 375:1100–1108.
4. Fleming DM (November 16, 2009) Influenza surveillance, the swine-flu pandemic, and the importance of virology. *Clin Evid*.
5. Rubin GJ, Potts HW, Michie S (2010) The impact of communications about swine flu (influenza A H1N1v) on public responses to the outbreak: Results from 36 national telephone surveys in the UK. *Health Technol Assess (Rocky)* 14:183–266.
6. Ferguson NM, et al. (2006) Strategies for mitigating an influenza pandemic. *Nature* 442:448–452.
7. Cooper BS, Pitman RJ, Edmunds WJ, Gay NJ (2006) Delaying the international speed of pandemic influenza. *PLoS Med* 3:e212.
8. Chowell G, Ammon CE, Hengartner NW, Hyman JM (2006) Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: Assessing the effects of hypothetical interventions. *J Theor Biol* 241:193–204.
9. Riley S, et al. (2003) Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science* 300:1961–1966.
10. Baguelin M, Van Hoek AJ, Flasche S, White PJ, Edmunds WJ (2010) Vaccination against pandemic influenza A/H1N1v in England: A real-time economic evaluation. *Vaccine* 28:2370–2384.
11. Bettencourt LMA, Ribeiro RM (2008) Real time Bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS One* 3:e2185.
12. Cauchemez S, Carrat F, Viboud C, Valleron AJ, Boëlle PY (2004) A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat Med* 23:3469–3487.
13. Fraser C, et al. (2009) Pandemic potential of a strain of influenza A(H1N1): Early findings. *Science* 324:1557–1561.
14. Bootsma MC, Ferguson NM (2007) The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proc Natl Acad Sci USA* 104:7588–7593.
15. Cauchemez S, Valleron A.-J, Boëlle P-Y, Flahault A, Ferguson NM (2008) Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* 452:750–754.
16. Wu JT, et al. (2010) School closure and mitigation of pandemic (H1N1) 2009, Hong Kong. *Emerg Infect Dis* 16:538–541.
17. Fleming DM, Zambon M, Bartelds AI (2000) Population estimates of persons presenting to general practitioners with influenza-like illness, 1987–96: A study of the demography of influenza-like illness in sentinel practice networks in England and Wales, and in The Netherlands. *Epidemiol Infect* 124:245–253.
18. Hall IM, Gani R, Hughes HE, Leach S (2007) Real-time epidemic forecasting for pandemic influenza. *Epidemiol Infect* 135:372–385.
19. Hardelid P, et al. (2010) Assessment of baseline age-specific antibody prevalence and incidence of infection to novel influenza AH1N1 2009. *Health Technol Assess (Rocky)* 14:115–192.
20. Cowling BJ, et al. (2010) Comparative epidemiology of pandemic and seasonal influenza A in households. *N Engl J Med* 362:2175–2184.
21. Bandanarayake D, Bissielo A, Huang S, Wood T (2010) Seroprevalence of the 2009 influenza A(H1N1) pandemic in New Zealand. (New Zealand Ministry of Health, Wellington) Technical report.
22. Flahault A, de Lamballerie X, Hanslik T, Salez N (2009) Symptomatic infections less frequent with H1N1pdm than with seasonal strains. *Plos Curr*, RRN1140, 10.1371/currents.RRN1140.
23. Carrat F, et al. (2008) Time lines of infection and disease in human influenza: A review of volunteer challenge studies. *Am J Epidemiol* 167:775–785.
24. Evans B, et al. (2011) Has estimation of numbers of cases of pandemic influenza H1N1 in England in 2009 provided a useful measure of the occurrence of disease? *Influenza Other Respi Viruses*, 5 10.1111/j.1750-2659.2011.00259.x.
25. Ong JBS, et al. (2010) Real-time epidemic monitoring and forecasting of H1N1-2009 using influenza-like illness from general practice and family doctor clinics in Singapore. *PloS One* 5:e10036.
26. Mossong J, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious disease. *PLoS Med* 5:e74.
27. Wallinga J, Lipsitch M (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc Lond B Biol Sci* 274:599–604.
28. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. *J Chem Phys* 21:1087–1091.
29. Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109.

Supporting Information

Birrell et al. 10.1073/pnas.1103002108

SI Text

1. SI Data. ***1.1. General practice (GP) consultation data.*** The United Kingdom (UK) Health Protection Agency (HPA) routinely utilizes GP consultation data for influenza-like illness (ILI) from two national GP surveillance networks. The Royal College of General Practitioners (RCGP) Weekly Returns Service is a sentinal GP network covering a weekly population of approximately 900,000 (1). The HPA/QSurveillance national surveillance system covers a weekly UK population of approximately 23 million. ILI data from both of these datasets are stratified by age and Strategic Health Authority (SHA). During the pandemic (April 2009—February 2010), the HPA used daily ILI reports that were routinely collected from the HPA/QSurveillance system to help in the public health reponse to the pandemic (2). GP consultation data were available from information provided for HPA surveillance during the pandemic.

1.2. Virological positivity. Both the RCGP and the HPA Regional Microbiology Network (RMN) (3) provide ongoing sentinel primary care surveillance with virological monitoring. Both networks cover populations of about 400,000 in England (RMN) and in England and Wales (RCGP) (4). This monitoring involves the taking of respiratory swabs from a subset of patients consulting GPs. A PCR assay is then employed to test the swabs for influenza strains as well as other respiratory virus infections. A more complete account of the virological monitoring undertaken by these two surveillance schemes can be found in ref. 5, but together they provide data on the positivity (the proportion of samples testing PCR positive) for a particular virus together with the epidemiological information attached to each sample. During the pandemic the schemes provided an estimate of virological positivity for influenza A/H1N1pdm in patients presenting ILI at GP consultations. Swab samples were included in the data only if the time between symptom onset and testing was 5 d or less, to avoid any adverse impact of test sensitivity.

1.3. Virologically confirmed cases. Management strategies over the initial stages of the pandemic were primarily concerned with the containment of the spread of the epidemic, prior to moving into a treatment phase. The initial containment phase was a period of enhanced surveillance during which contacts of known infected individuals were traced and laboratory confirmations of the infection were obtained wherever possible. This work resulted in the generation of the FF100 and the FluZone databases (5, 6). Routine laboratory confirmations were discontinued on June 25, but we use here the data only up to June 19 to allow for the gradual cessation in the collection of this type of data.

1.4. Serological studies. The HPA runs an annual collection of serum samples for the purpose of carrying out cross-sectional antibody prevalence studies (7). Collected sera are residual samples from those submitted to microbiological laboratories for antibody testing or diagnostic screening. During the 2009 pandemic these were supplemented by further samples from biochemistry laboratories. Testing of samples collected in 2008 provided an age-specific estimate of baseline prevalence of antibodies to the A/H1N1pdm virus, giving an indication of the level of preexisting immunity in the population. Among this baseline sample, a haemagglutin-inhibiting antibody titer of 32 is sufficient to indicate protection against A/H1N1pdm influenza in approximately 50% of individuals (8–10), although protection is often present in individuals with titer values ≥8. Despite this uncertainty in the level of protection, assays ≥32 are assumed to be protective against A/H1N1pdm, though this may well be an underestimate of the preexisting immunity. During the A/H1N1pdm epidemic, further samples were submitted monthly to the HPA and tested similarly. These data have been published elsewhere (11). A geometric-mean titer value of ≥32 is assumed to arise as a result of recent infection, so that individuals with this titer value can be assumed to either belong to the group "protected" at baseline or to have been recently infected. A 2-wk delay between infection and seroconversion is assumed, so each individual sample is assumed to be representative of the level of infection among the population 14 d prior to the sampling date.

1.5. Further information. 1.5.1. National Pandemic Flu Service (NPFS). On July 23, the NPFS phone and Internet service was launched. This system was designed to relieve the pressure put upon the general practice network by allowing an online assessment of the suitability of patients for an antiviral prescription. If approved, patients were given a unique code that they could use to collect antivirals drugs.

1.5.2 Flusurvey. This was a web-based resource (12) for registered users to report, on a weekly basis, their experience of ILI over the course of the preceding week. The participants report any symptoms experienced and, if applicable, their subsequent health-care-seeking behavior. Participants register on an entirely voluntary basis.

In addition to the data described in the preceding subsections, information on population sizes by age group were taken from the 2008 midyear population estimates made by the Office of National Statistics (13).

2. SI Materials and Methods. ***2.1. Model details. 2.1.1. The transmission model.*** The spread of the disease among the population is modeled using an SEIR transmission model, comprising states S (individuals susceptible to infection), E (individuals with latent infection), I (infectious individuals), and R (individuals no longer at risk of acquiring or transmitting infection). The E and I compartments are further subdivided into two substates E_1, E_2 and I_1, I_2, respectively), leading to waiting times in the E and I states, which are distributed according to gamma distributions.

The transmission model (as can be seen in Fig. 2 of the main article) is governed by a time- and age-varying force of infection, $\lambda(t,a)$, and transition rate parameters σ and γ. These transition rates are related to the mean latent period, d_L, and the mean infectious period, d_I, by

$$\sigma = 2/d_L, \qquad \gamma = 2/d_I.$$

The model dynamics arise as a discretization of a system of differential equations. Assuming δt small relative to the expected waiting times, σ^{-1}, γ^{-1} the difference equations corresponding to the nth time step are

$$
\begin{aligned}
S(t_n,a) &= S(t_{n-1},a)(1-p_\lambda(t_{n-1},a))\\
E_1(t_n,a) &= E_1(t_{n-1},a)(1-\sigma\delta t) + S(t_{n-1},a)p_\lambda(t_{n-1},a)\\
E_2(t_n,a) &= E_2(t_{n-1},a)(1-\sigma\delta t) + E_1(t_{n-1},a)\sigma\delta t\\
I_1(t_n,a) &= I_1(t_{n-1},a)(1-\gamma\delta t) + E_2(t_{n-1},a)\sigma\delta t\\
I_2(t_n,a) &= I_2(t_{n-1},a)(1-\gamma\delta t) + I_1(t_{n-1},a)\gamma\delta t\\
R(t_n,a) &= R(t_{n-1},a) + I_2(t_{n-1},a)\gamma\delta t,
\end{aligned}
\qquad \text{[S1]}
$$

where $t_n = n\delta t$ and $p_\lambda(t_n,a)$ is the proportion of susceptibles infected in the interval $[t_{n-1},t_n)$. $p_\lambda(t_n,a)$ is closely linked to the force of infection and is dependent on two key quantities, the basic reproduction number of the virus, R_0, and a time-dependent mixing matrix $\boldsymbol{M}(t)$, giving the relative rates of contact between individuals in different population strata, which in this case are defined by the age groupings (seven groups: <1, 1–4, 5–14, 15–24, 25–44, 45–64, 65+). We relate these quantities through the Reed–Frost formulation for the infection hazard (14), so that

$$
\begin{aligned}
p_\lambda(t_n,a) &= \left(1-\prod_{b=1}^{7}\left[(1-p_\beta(t_n)_{a,b}R_0\gamma/2)^{I_1(t_n,b)+I_2(t_n,b)}\right]\right)\delta t\\
&\approx \lambda(t_n,a)\delta t.
\end{aligned}
\qquad \text{[S2]}
$$

The time-dependent matrix

$$
\boldsymbol{p}_\beta(t) = \{p_\beta(t)_{a,b}\} = \boldsymbol{M}(t)/R_0^* \qquad \text{[S3]}
$$

is a scaled version of the mixing matrix, with R_0^* being the dominant eigenvalue of the scaled next generation matrix $\boldsymbol{M}^*$, whose (a,b)th element is given by $M^*_{a,b} = M(0)_{a,b}N_a$, where N_a is the population size within age group a.

If t_0 is selected to be sufficiently close to the emergence of the epidemic within the study population, then we can use as an initial condition that the occupancies of disease states E_1, E_2, I_1, and I_2 are growing exponentially with rate ψ_r. This leads to the time-zero solution:

$$
\begin{aligned}
I_1(0,a) &= I_a^{\text{tot}}\left(1+\frac{\epsilon\gamma\delta t}{\alpha+\gamma\delta t}\right)^{-1} \qquad I_2(0,a) = (I_a^{\text{tot}} - I_1(0,a))\\
E_2(0,a) &= I_1(0,a)\frac{\alpha+\gamma\delta t}{\sigma\delta t} \qquad E_1(0,a) = E_2(0,a)\frac{\alpha+\sigma\delta t}{\sigma\delta t},
\end{aligned}
$$

where $\alpha = \exp(\psi_r\delta t) - 1$ and $\boldsymbol{I}^{\text{tot}}$ is the vector of the initial number of infectious individuals in each age group. Prior to the onset of the influenza epidemic, there is some preexisting immunity within each population. If we denote the proportion of the population within age group a who have a prior immunity to A/H1N1pdm by $1-\pi_a$, then we set $R(0,a) = (1-\pi_a)N_a$, and $S(0,a) = N_a - E_1(0,1) - E_2(0,a) - I_1(0,a) - I_2(0,a)$.

In a development of earlier work (15), the value of the reproduction number during the initial exponential growth period has been found to be given by the relation (16)

$$
R_0^{\text{init}} = \psi_r d_I \frac{(\frac{\psi_r d_L}{2}+1)^2}{1-\frac{1}{(\frac{\psi_r d_I}{2}+1)^2}}. \qquad \text{[S4]}
$$

Rather than attempting to estimate the initial infectious population size in each of the age groups, we can instead observe that if we define a quantity β to be the average daily rate of secondary infections per primary infection, then the average initial infection hazard is given by $\lambda_0 = \beta\sum_a I_a^{\text{tot}}$. Similarly, we can relate β to the reproduction number at time 0 through $R_0^{\text{init}} = \beta\sum_a N_a d_I$. Equating these expressions for β gives

$$
I_+ = \sum_a I_a^{\text{tot}} = \frac{d_I\lambda_0\sum_a N_a}{R_0^{\text{init}}}. \qquad \text{[S5]}
$$

The individual I_a^{tot} are then determined by multiplying this sum by the eigenvector corresponding to the dominant eigenvalue of the matrix $M(0)$, scaled so that the vector components sum to 1. If this eigenvector is χ^M,

$$
I_a^{\text{tot}} = I_+\chi_a^M. \qquad \text{[S6]}
$$

2.1.2. Disease severity and disease reporting. The disease and reporting models (see Fig. 2 of the main article) take as input the modeled time series of the number of new infections by age. From the block of Eq. **S1**, this can be seen to be

$$
\Delta_{\text{infec}}(t_n,a) = S(t_{n-1},a)p_\lambda(t_{n-1},a). \qquad \text{[S7]}
$$

A significant proportion of these new infections does not manifest in the form of febrile symptoms. As a result infections go undetected and will not appear in surveillance data. If θ represents the proportion of infections with associated symptoms, the number of incident symptomatic cases at time t within age group a is given by

$$
\Delta(t_n,a) = \theta\Delta_{\text{infec}}(t_n,a).
$$

This quantity we refer to here onward as the incident number of cases.

It is assumed that each case will be reported as present in one of the datasets with some probability, independently of whether or not they contribute to any of the other datasets. If we focus on the GP consultations, we define a time- and age-varying propensity to consult as a result of symptoms to be $p_{\text{GP}}(t,a)$. Given this, we then need to consider the delay between the time of infection and the time of reporting. Such delay is made up of three components: the incubation period, i.e., the time from infection to symptoms; the waiting time from symptom onset to the healthcare event; and the time taken from the event to its appearance in data, i.e., the reporting delay.

If we consider each of the intermediate steps, we have that the expected number of reported events at time t in age group a can be expressed as

$$
\begin{aligned}
\mu_{t_n a} = \sum_{k=0}^{n} & f_{(\mu_{\text{incubation}},\sigma^2_{\text{incubation}})}(k,a)\sum_{l=0}^{n-k} f_{(\mu_{\text{event}},\sigma^2_{\text{event}})}(l,a)p_{\text{event}}(t_{n-k-l},a)\\
&\times\sum_{m=0}^{n-k-l} f_{(\mu_{\text{report}},\sigma^2_{\text{report}})}(m,a)\Delta(t_{n-k-l-m},a),
\end{aligned}
\qquad \text{[S8]}
$$

where the $f_{(\mu_\cdot,\sigma^2)}(t,a)$ give the probability that a gamma random variable indexed by mean $\mu_\cdot$ and variance σ^2 lies in the interval $(t,t+1)\delta t$. Convolution over three indices is computationally very expensive and it greatly diminishes our power to run a model such as this in real-time. Here we approximate the compounded waiting time (the sum of the incubation time, time to event, and reporting times) by a single gamma distribution, the mean and variance of which are the sums of the component means and of the component variances, respectively (17). Using such an approximation, Eq. **S8** reduces to a single summation

$$
\begin{aligned}
\mu_{t_n a} = & p_{\text{event}}(t_n,a)\\
&\times\sum_{k=0}^{n}\Delta(t_{n-k},a)f_{(\mu_{\text{incubation}}+\mu_{\text{event}}+\mu_{\text{report}},\sigma^2_{\text{incubation}}+\sigma^2_{\text{event}}+\sigma^2_{\text{report}})}(k,a).
\end{aligned}
\qquad \text{[S9]}
$$

Even with this approximation, it is this convolution that is responsible for the majority of the computational expense.

2.2. Interventions and modeling challenges. The spread of a novel influenza virus leads to changes in the way the public interacts both with each other and with the health services. In the initial stages of this epidemic, it is possible that a heightened state of alertness among the public induced a surge in the number of people presenting ILI symptoms to their GPs. In turn, throughout the treatment phase of the epidemic, government advice on the management and reporting of ILI symptoms changed as the NPFS was launched. Later on in the epidemic, the initiation of a vaccination program once again changed the way the public interacts with the health services.

We address these issues by considering relevant model parameters to be piecewise-constant over time, with change points positioned as close as possible to the behavioral change-causing event.

2.2.1. School holidays. Although the widespread closure of schools as a measure to mitigate the spread of the emergent epidemic was never enforced, the epidemic did straddle a number of school holidays including the six-week summer holiday common to all English schools. As children younger than 15 y old are key agents of transmission, it is not surprising that this led to some dying down of the epidemic in these periods. School-holiday effects upon transmission can be incorporated via the mixing matrices $\boldsymbol{M}(t)$.

Initially, mixing matrices obtained from the POLYMOD (Improving Public Health Policy in Europe through Modeling and Economic Evaluation of Interventions for the Control of Infectious Diseases) study into social mixing patterns (18) were employed. This study, through presenting estimated rates of physical contact between members of the different age groups both on weekends and weekdays, provides "holiday" and "non-holiday" contact matrices to be used for periods corresponding to without and within school terms, respectively. However, these matrices were found to underplay the role of the <15 y-old age groups. This arose due to the near-simultaneous introduction of the NPFS and start of the school summer holidays. The falloff in GP consultation could therefore be attributable to either a mirrored fall in the propensity to consult or through contact patterns. Using the POLYMOD matrices alone led to an estimated drop in the $p_{\mathrm{GP}}(t,a)$ to values orders of magnitude smaller than observed in Flusurvey (12). Parameterizing these matrices to allow for both estimation of the effect of school holidays on contact rates and the relative likelihood of transmission in infectious contacts involving adults compared to those that just involve children (see also refs. 19 and 20) leads to more plausible estimates.

Specifically, this parameterization is achieved by considering a base matrix that we take to be the nonholiday matrix from the POLYMOD study. Denoted by M_{POLY}, the (i,j)th element of this matrix gives the rate of contact between a member of the population in age group i with someone in age group j. To move from this matrix to $\boldsymbol{M}(t)$ of Eq. **S3**, consider the matrices $\boldsymbol{A}_{\mathrm{base}}$, $\boldsymbol{A}_{\mathrm{summer}}$, $\boldsymbol{A}_{\text{half-term}}$ given respectively in Eqs. **S10**–**S12**:

$$\boldsymbol{A}_{\mathrm{base}} = \begin{pmatrix} 1 & 1 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & 1 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & 1 & 1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \end{pmatrix}. \qquad \textbf{[S10]}$$

$$\boldsymbol{A}_{\mathrm{summer}} = \begin{pmatrix} 1 & 1 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & m_2 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & 1 & m_3 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \end{pmatrix}, \qquad \textbf{[S11]}$$

$$\boldsymbol{A}_{\text{half-term}} = \begin{pmatrix} 1 & 1 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & m_4 & 1 & m_1 & m_1 & m_1 & m_1 \\ 1 & 1 & m_5 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \\ m_1 & m_1 & m_1 & m_1 & m_1 & m_1 & m_1 \end{pmatrix}. \qquad \textbf{[S12]}$$

Using the $\odot$ symbol to indicate componentwise multiplication of matrices such that $\boldsymbol{A} = \boldsymbol{B} \odot \boldsymbol{C} \Rightarrow A_{ij} = B_{ij}C_{ij}$, $\boldsymbol{M}(t)$ is then assumed to be

$$\boldsymbol{M}(t) = \begin{cases} \boldsymbol{A}_{\mathrm{summer}} \odot \boldsymbol{M}_{\mathrm{POLY}} & t \in (79,128] \\ \boldsymbol{A}_{\text{half-term}} \odot \boldsymbol{M}_{\mathrm{POLY}} & t \in (23,30] \cup (177,184] \cup (233,245] \\ \boldsymbol{A}_{\mathrm{base}} \odot \boldsymbol{M}_{\mathrm{POLY}} & \text{otherwise.} \end{cases} \qquad \textbf{[S13]}$$

The time intervals (23,30] days and (177,184] days correspond to the half-term school holidays in May and October, respectively, and these holidays share parameters with the interval (233,245], which corresponds to the Christmas holiday. The interval (79,128] days corresponds to the longer summer school holiday. Parameter m_1 features in all matrix entries corresponding to contact rates involving an adult age group (i.e., ≥15 y), conferring upon it the interpretation as the proportionate reduction in the probability of an infectious contact involving an adult being "effective" (i.e., leading to disease transmission) in comparison to an infectious contact involving two children under the age of 15 y.

Correspondingly, parameters m_2 and m_4 give the proportionate reduction in the contact rates among the 1–4 y age group in the summer and half-term holidays, respectively, with parameters m_3 and m_5 similarly defined for the 5–14 age group. The two holiday periods are treated differently, through the assignment of distinct parameters to allow for the possibility that people behave differently in these periods. During the summer holiday it is expected that a larger proportion of the children will spend time overseas or on some sort of prolonged vacation.

2.2.2. Changing public behavior (1): Acting on government advice. As already discussed, it cannot be assumed that the propensity of those with symptoms of A/H1N1pdm to consult with a GP is constant within each age group over time. In particular, the issue of modeling the propensity of patients to consult with a GP is complicated further by the introduction of the NPFS on July 23. The natural effect of the NPFS was to drastically reduce the proportion of individuals who appear in the GP consultation datasets (pregnant women, children <1 y, and anyone else occupying a particular high-risk group were still advised to consult directly with a GP). To account for this we parameterize $p_{\mathrm{GP}}(t,a)$ such that children and adults have distinct propensities to consult at all times and these are subject to two change points (see Fig. S1), the first of which is chosen to correspond to the introduction of NPFS. The second breakpoint is designed to bisect the interval between the introduction of the NPFS and the end of the dataset. Its inclusion is to allow for the possibility that the advice for patients to make the NPFS their first port of call was perhaps being

forgotten over time as people reverted to their traditional health-seeking behavior (i.e., going to the GP).

Specifically, we parameterize $p_{\mathrm{GP}}(t,a)$ as follows:

$$p_{\mathrm{GP}}(t,a) = \begin{cases} p_0 & a < 15\text{ y}, t \le 83\text{ d} \\ p_1 & a \ge 15\text{ y}, t \le 83\text{ d} \\ p_0 p_2 & a < 15\text{ y}, 83 < t \le 137\text{ d} \\ p_1 p_3 & a \ge 15\text{ y}, 83 < t \le 137\text{ d} \\ p_0 p_4 & a < 15\text{ y}, 137 < t \le 192\text{ d} \\ p_1 p_5 & a \ge 15\text{ y}, 137 < t \le 192\text{ d} \end{cases}, \qquad \textbf{[S14]}$$

where day 84 (July 23) corresponds to the introduction of the NPFS phone line and day 137 (September 14) and day 192 (November 8) form an equal partition of the time from the NPFS advent until the end of data currently under analysis. By parameterizing in this way and constraining the p_i to take values in the range [0,1], we guarantee that the propensity to consult is at a peak in the period prior to the NPFS introduction.

2.2.3. Changing public behavior (2): The worried well. In addition to the change in the health-care-seeking behavior of infected individuals with flu-like symptoms, the change in behavior of individuals suffering from symptoms derived from non-A/H1N1pdm ILI also needs to be considered. These individuals also appear in the GP consultation data, contributing a time-varying level of noise.

The data on virological positivity allow us to partition the observed ILI consultations among those that are genuine pandemic A/H1N1pdm cases and those that are not. This is done externally to our integrated model of Fig. 2 (in the main article) by modeling collateral data on positivity and GP consultation rates from other regions of England.

Joint regression models were fitted to the consultation rate and positivity data from the regions outside of London. Only regions outside of London were used, to avoid making double use of the GP consultation and virological positivity data. This joint model provides posterior means and covariance matrices for the parameters of a log-linear regression of the background consultation rate that assumes linearly independent temporal and age effects and an interaction that takes the form of a change point in the age effects at the time of NPFS introduction:

$$\log(B(t,a)) = \alpha_t + \beta_a + \mathbb{1}_{\{t>t_{\mathrm{NPFS}}\}}\beta'_a.$$

This regression equation is imported into the integrated model, together with a multivariate normal prior distribution for $(\alpha_t,\beta_a,\beta'_a)$, which conserves the posterior means and covariances found through the external model. This allows estimation of rates of background consultation in London, which are influenced by the equivalent rates from the other regions of England.

2.3. Inference. 2.3.1. Likelihood. The GP consultation data are assumed to have a negative-binomial distribution. We parameterize the negative-binomial distribution in terms of a mean number of consultations, μ_{ta} (as found from Eq. **S9**) and a piecewise-constant dispersion parameter, $\eta_t > 1$, such that the likelihood component is

$$L(y_{ta}|\varphi) = \frac{\Gamma(y_{ta}+r_{ta})}{\Gamma(r_{ta})\Gamma(y_{ta}+1)}\left(\frac{1}{\eta_t}\right)^{r_{ta}}\left(1-\frac{1}{\eta_t}\right)^{y_{ta}} \qquad \eta_t \ge 1,$$

where $r_{ta} = \mu_{ta}/(\eta_t - 1)$. The dispersion parameter η_t has one change point at the time of the NPFS launch, reflecting a fundamental modification to the surveillance system. With this parameterization, $\mathbb{E}Y(t,a) = \mu_{ta}$ and $\mathrm{Var}(Y(t,a)) = \mu_{ta}\eta_t$.

2.3.2. Priors. Table S1 provides a near comprehensive list of the parameters that comprise the vector φ, together with their assumed fixed value or prior distribution. Parameters of the background model are omitted for brevity. The chosen time step for the model is $\delta t = 1/2$ d, which limits how small both d_I and d_L can be without invalidating approximations made in the discretization of the model dynamics. Therefore the mean infectious period is estimated over values constrained to be at least 2 d, whereas we follow ref. 19 in fixing the mean latent period to 2 d. The proportions susceptible were determined from the estimated baseline prevalence of A/H1N1pdm antibodies (11). The choice of prior for θ is based on data reported in ref. 21. The parameters of the incubation period distribution were taken from recent work on confirmed cases with identifiable exposure times (22). For the parameters governing the propensity to consult with a GP, p_i, priors were derived from (nonage specific) estimates for the United Kingdom available from the Flusurvey study (12), with some appropriate statements of uncertainty attached.

3. SI Further Results. ***3.1. Goodness of fit.*** Fig. S2 compares the predicted age-group specific and weekly aggregated GP consultations with the observed data. The predicted consultations track the data rather well, with perhaps some overestimation of consultations during the initial epidemic growth and some underestimation around the peak of the second wave. Fig. S3 features the corresponding plot for the virological positivity data. Here, the model predicted positivity is plotted together with a 95% credible interval. The observed positivities are overlaid and should differ from the modeled positivities through binomial error. The model appears to perform well, albeit with some underestimation of the positivity in the 15–24 age group, perhaps due to the influence of the 25–44 age group, the data for which had far larger sample sizes. Of course, there is some arbitrariness introduced via the choice of age groups and any lack of fit within the 15–24 age group could be due to heterogeneity within this age group arising from the fact that this group includes a significant proportion of individuals who are still of school age. Similar within-group heterogeneity might explain the lack of fit to the serology data within this age group evident in Fig. S4, although this could also be explained through underestimation of the baseline levels of immunity within the age group. Elsewhere, the serological data seem highly plausible given the model.

3.2. Sensitivity analyses. 3.2.1. An imperfect virological test. The proportion, $q_{t_n a}$, of GP consultations that are attributable to A/H1N1pdm at the nth time step, t_n, and within any given age group, a, is given by

$$q_{t_n a} = \frac{\mu_{t_n a}}{\mu_{t_n a} + B(t_n,a)}.$$

Thus far, this proportion has been assumed to be unbiasedly informed by the virological swabbing data, with these data, $W_{t_n a}$, taken to be distributed

$$W_{t_n a} \sim \mathrm{Bin}(M_{t_n a}, q_{t_n a}).$$

It is, however, more likely that the testing procedure is imperfect, i.e., with sensitivity less than 1. This may be due to the test itself, or perhaps, in some individuals, to the depletion of the viral load to an undetectable level over the interval from infection to swabbing (23).

Therefore, what we see is an apparent positivity, $q^*_{t_n a}$, which relates to the test sensitivity, k_{sens}, via the equation $q^*_{t_n a} = k_{\mathrm{sens}} q_{t_n a}$, assuming that specificity is equal to 1. The positivity data are assumed

$$W_{t_n a} \sim \mathrm{Bin}(M_{t_n a}, q^*_{t_n a}).$$

Intuitively, therefore, an imperfect test would lead us to expect that there is a greater number of GP consultations for A/H1N1pdm in the population than would be found through naive multiplication of the observed consultations with the observed positivity.

Table S2 displays estimates based on posterior medians (and credible intervals) for some epidemic quantities under the assumption of a perfect test, under two reasonable guesses for the sensitivity ($k_{\text{sens}} = 0.9,0.8$) and an extreme case ($k_{\text{sens}} = 0.5$). Counterintuitively, the overall incidence actually decreases with decreasing test sensitivity, along with the average reproductive number, R_0. Specifically, this decrease occurs in the number of infections in the second wave, as we can see that the posterior probability that the first wave has a higher peak rate of infection increases with decreasing test sensitivity. However, a decreasing infection attack rate (IAR) is accompanied by an increasing symptomatic attack rate (SAR). This arises through the proportion symptomatic increasing with decreasing test sensitivity, which more than compensates for the drop in infection rates when fitting to the GP consultation data. The model can explain the "true" positivity, $q_{t_n a}$, being larger than the observed positivity, $q^*_{t_n a}$, through either (*i*) higher levels of overall infection or (*ii*) a greater proportion of the infections appearing in data, due to a higher degree of symptomatic infection or to increased rates of reporting. Here, case (*ii*) appears to be occurring.

Of the four examples here, the posterior distribution for the likelihood takes larger values in the case $k_{\text{sens}} = 0.8$, with the extreme case performing much worse than the other three cases. Despite this, inferences do not seem to be substantially altered over the three "feasible" values of k_{sens}, so we consider them to be robust to the assumption of a perfect test.

3.2.2. Contact patterns. The symmetric matrix of mixing rates, $\boldsymbol{M}(t)$ (cf. Eq. **S3**) is populated with values taken from the UK POLYMOD data, based on all types of contacts on weekdays, as done elsewhere (24). Even the more flexible parameterization of these matrices adopted here (Eq. **S13**) constitutes a strong set of restrictive assumptions. For example, what happens if the estimated POLYMOD matrices are inadequate or inappropriate for describing the patterns of infectious contacts? Furthermore, in the main text, Table 2 highlights the difficulty in estimating school-holiday effects among the 1–4 y-old age group (the posterior distribution looks very similar to the prior). If we adopt a more parsimonious parameterization, by assuming the impact of the school holidays upon the social dynamics within the 1–4 age group is identical to the impact in the 5–14 age group ($m_2 = m_3$ and $m_4 = m_5$), do we obtain substantially different inferences?

To examine these questions, we repeated the modeling exercise using instead the contact matrix derived from the UK subset of the POLYMOD data on weekend only contacts. This proved to have only an imperceptibly small effect on key estimates. Alternatively, if the constraints $m_2 = m_3$ and $m_4 = m_5$ are enforced and denoted m^*_3 and m^*_5, respectively, our estimated posterior medians and credible intervals are

$$m_1 = 0.468(0.405,0.539)$$
$$m^*_3 = 0.336(0.128,0.485)$$
$$m^*_5 = 0.483(0.249,0.755).$$

The posterior summaries presented here for m^*_3 and m^*_5 resemble those presented for parameters m_3 and m_5 in Table 2 (in the main text), with some slight movement toward the estimates obtained for parameters m_2 and m_4 in the main analysis. Once again, estimates of other parameters and epidemic quantities are robust to the parameterization choice. Although this choice is slightly more parsimonious than the parameterization in the main text, with no significant loss of model fit, we prefer the main text analysis, as it is a strong a priori belief that the two age groups should be treated differently (1 y-olds and 14 y-olds will be affected in very different ways by breaks in school terms) and, again, the key inferences are left unaltered.

As inference appears robust to the precise choice of the values of the matrices $\boldsymbol{M}(t)$, we further investigate the feasibility of actually estimating these values. The contact matrices are symmetric and therefore only have free entries in the upper-triangular portion of the matrix. Furthermore, $p_\lambda(t,a)$, as defined in Eq. **S2**, is independent of the scale of the mixing matrices for all values of t and a. It follows that the likelihood is similarly independent. Therefore, for identifiability, one matrix entry is assigned a fixed value, with other matrix entries corresponding to rates of mixing relative to the fixed entry. So for the 7×7 mixing matrix, we have 27 free parameters. As this is an illustrative exercise only, a bootstrap sample of contact matrices was drawn based on the UK POLYMOD data using the approach of ref. 25. Each matrix is then normalized so that the M_{33} entry is equal to 1. It is no longer possible to identify a parameter with the same interpretation as parameter m_1, but, to factor in different rates of infectivity between adults and children, each bootstrap matrix was assigned a realization of a $U(0,1)$ random variable, and the entries corresponding to mixing rates involving adults were multiplied by this value. Based on this bootstrap sample, the means and variances for each of the 27 free matrix entries were used as the basis for gamma priors. Again, holiday effects were incorporated as previously, through parameters $m_2,\ldots,m_5$.

Fig. S5 plots the posterior distributions for these parameters in the form of an upper-triangular matrix. It can be seen that for many of the mixing rates, the posterior distributions are similar to the prior distributions, particularly for those parameters corresponding to at least one of the <1 and the >65 age groups. In the remaining age groups, the prior to posterior shift would appear to be toward lower values. As these rates are all measured relative to the M_{33} entry, this suggests that the POLYMOD matrices alone, even accounting for infectivity could, potentially, be underplaying the importance of disease transmission among the 5- to 14-y-old age group. Only one matrix element is inflated in moving from prior to posterior, M_{34}, giving relative rates of effective contact between the 5–14 and 15–24 age groups. However, these contacts still only occur at a posterior median of 0.272 times the rate of the within 5- to 14-y-old contacts. The second column of Table S3, which gives the estimated values for R_0 and the first and second wave attack rates, shows a higher level of total incidence in these results than in the analysis of the main text, despite a much lower estimate of R_0. As the densities of Fig. S5 would appear to suggest, the higher total incidence is due to much higher levels of incidence in the 15–24 age group, which, compared to the child-age groups is highly populated with a supply of susceptibles that is not so easily exhausted.

This component of the sensitivity analysis is presented here as a proof-of-principle. The choice of prior distribution is key here, due to the influence it exerts on the pattern of incidence and further study is ongoing to develop this in greater detail.

3.2.3. The propensity to consult. The piecewise-constant parameterization of $p_{\text{GP}}(t,a)$ used to obtain the results presented in the main text was one that developed over the course of the pandemic out of the necessity to maintain an adequate fit to the GP consultation data. However, the step changes involved might be considered a crude approximation to the underlying process. As a test of the sensitivity, we try two alternate models for $p_{\text{GP}}(t,a)$. Using the same breakpoints, we consider $p_{\text{GP}}(t,a)$ to be piecewise-linear over time on the logistic scale, either from $t = t_{\text{NPFS}}$ onward (model 1) or at all times (model 2). These two models require the addition of two and four new parameters, respec-

tively. Fig. S6 plots the estimated temporal evolution in the propensity to consult for the two models for the adult age groups. Post-NPFS the two plots look very similar and differ from the piecewise-constant version in that there is a decline in the consultation propensity in the interval following the day-178 breakpoint, suggesting that the consultation propensity might be mirroring the pattern of consultations rather than gradually restoring itself to some pre-A/H1N1pdm level. In the pre-NPFS interval, the influence of the peak consultation rate is evident, given that the constant propensity in model 1 is equal to the propensity in model 2 at $t = t_{\mathrm{NPFS}}$, rather than being an average value. In model 2, the time-0 value of the propensity appears to be highly prior influenced and so there is little learning to be made here about the propensity to consult over this interval. Because of this prior influence, in model 2, the proportion symptomatic θ is estimated to be lower than in model 2, to counteract this increased propensity to consult.

3.3. Further applicability. As has been discussed in *SI Data*, the data in England was subdivided into 10 distinct regions defined by the boundaries of the SHAs. The highly localized collection of virological and serological data is such that sample sizes in each region were too small to enable accurate estimation of the epidemic in many SHAs, whereas, if monitoring at the national level, the data are unrepresentative.

Taking this into consideration, we repeated our modeling methodology for one other SHA, the West Midlands, and two further aggregated regions each consisting of four of the remaining eight SHAs, "North" and "South." Fig. S7 plots the estimated epidemic curve for each region, whereas Table S4 contains posterior estimates for key epidemic quantities.

Fig. S7 shows that, for the two epidemics that experienced significant levels of first wave infection, London and the West Midlands, their first wave peaks occur simultaneously and at a similar height. This constitutes a depletion in the population susceptibility of the West Midlands (it has a much smaller population) that is reflected in the lower levels of incidence in the second wave. Such a distinctive two-wave pattern of infection was not obvious in the North and South, however. The South shows some first wave activity, but the real peak in incidence rates is during the second wave. The North behaves differently again. The lack of any significant first wave of infection makes the school-holiday effects impossible to estimate—how can a reduction in incidence be estimated when there is negligible incidence?—and thus all infection takes place in one big autumn wave, although a considerable level of uncertainty is attached, possibly, in part, arising from the very low virological sample sizes in the North.

The estimates of θ in Table S4 remain reasonably consistent across all four regions; each credible interval has a large intersection with the others. London and West Midlands display comparable estimates for R_0, whereas the North and South have much lower values. It is speculated that this may well arise as a result of the different levels of population density: The London and West Midlands regions encompass the two biggest cities in the country. Again, London and West Midlands achieve similar estimates for parameter m_1, whereas this parameter is close to 1 for the North and South. The age profile of infected individuals is clearly more uniform in the North and South regions, which may have something to do with the bulk of infection occurring in the autumn and winter months, when older age groups are at greater risk. School-holiday effects are generally estimated with wide credible intervals and as a result, the effect of the half-term holiday (m_5) as estimated in the West Midlands would appear to be the only anomalous result here.

3.4. Serology. In the main text, *Predicting the Epidemic*, we show how the model's ability to make convergent epidemic estimates coincides with the introduction of some serological survey data. To further understand the role and the importance of these data, and to test whether this is mere coincidence, we repeated the modeling exercise of the main text, but with all serological data removed. The 83-, 143-, 192-, and 245-d estimates are presented in Fig. S8. Here, it is only the 245-d analysis that looks like a feasible representation of the epidemic. This indicates that, although post hoc epidemic reconstruction may well be feasible, the timely availability of serological data is a requisite for real-time inference.

1. Fleming DM (1999) Weekly Returns Service of the Royal College of General Practitioners. *Commun Dis Public Health* 2:96–100.
2. Harcourt SE, et al. (2011) Use of a large general practice syndromic surveillance system to monitor the progress of the influenza A(H1N1) pandemic 2009 in the UK. *Epidemiol Infect*, 10.1017/S095026881100046X.
3. McCartney C (2008) Regional microbiology network. *Br J Infect Control* 8:28–29.
4. Health Protection Agency (2009) HPA surveillance systems: Swine influenza (H1N1swl). http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1243467944074.
5. Health Protection Agency (2010) Epidemiological report of pandemic (H1N1) 2009 in the UK. http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1284475321350.
6. Health Protection Agency (2009) Positivity rates from virological sampling. http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1259152231549.
7. Osborne K, Gay N, Hesketh L, Morgan-Capner P, Miller E. (2000) Ten years of serological surveillance in England and Wales: Methods, results, implications and action. *Int J Epidemiol* 29:362–368.
8. Al-Khayatt R, Jennings R, Potter CW (1984) Interpretation of responses and protective levels of antibody against attenuated influenza A viruses using single radial haemolysis. *J Hyg (Lond)* 93:301–312.
9. Hobson D, Curry RL, Beare AS, Ward-Gardner A. (1972) The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses. *J Hyg (Lond)* 70:767–777.
10. de Jong JC, et al. (2003) Haemagglutination-inhibiting antibody to influenza virus. *Dev Biol (Basel)* 115:63–73.
11. Miller E., et al. (2010) Incidence of 2009 pandemic influenza a h1n1 infection in England: A cross-sectional serological study. *Lancet* 375:1100–1108.
12. Flusurvey. (2009) http://www.flusurvey.org.uk.
13. Office for National Statistics. (2010) Super output area mid-year population estimates for England and Wales (experimental), http://data.gov.uk/dataset/super_output_area_mid-year_population_estimates_for_england_and_wales_experimental_-_mid-2008.
14. Ball F. (1983) A threshold theorem for the Reed-Frost chain-binomial epidemic. *J Appl Probab* 20:153–157.
15. Anderson D, Watson R. (1980) On the spread of a disease with gamma distributed latent and infectious periods. *Biometrika* 67:191–198.
16. Wearing HJ, Rohani P, Keeling MJ (2005) Appropriate models for the management of infectious disease. *PLoS Med* 2:e174.
17. Stewart T, Strijbosch L, Moors J, Batenburg PV (2007) A simple approximation to the convolution of gamma distributions (revision of dp 2006-27), (Tilburg University, Center for Economic Research), Discussion Paper 2007-70.
18. Mossong J, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious disease. *PLoS Med* 5:e74.
19. Vynnycky E, Edmunds WJ (2008) Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures. *Epidemiol Infect* 136:166–79.
20. Wu JT, et al. (2010) School closure and mitigation of pandemic (H1N1) 2009, Hong Kong. *Emerg Infect Dis* 16:538–541.
21. Monto AS, Koopman JS, Longini IM (1985) Tecumseh study of illness. XIII. Influenza infection and disease, 1976–1981. *Am J Epidemiol* 121:811–822.
22. Tom BDM, et al. (2011) Estimating time to onset of swine influenza symptoms after initial novel A(H1N1v) viral infection. *Epidemiol Infect* 139:1418–1424.
23. Fleming DM (2009) Influenza surveillance, the swine-flu pandemic, and the importance of virology. *Clin Evid* November 16, 2009.
24. Baguelin M, Van Hoek AJ, Flasche S, White PJ, Edmunds WJ (2010) Vaccination against pandemic influenza A/H1N1v in England: A real-time economic evaluation. *Vaccine* 28:2370–2384.
25. Hens N, et al. (2009) Estimating the impact of school closure on social mixing behaviour and the transmission of close contact infections in eight European countries. *BMC Infect Dis* 9:187+.

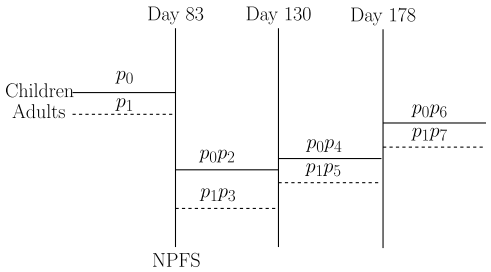

Fig. S1. Propensities to consult a GP over time given symptoms of infection with the A/H1N1pdm virus.

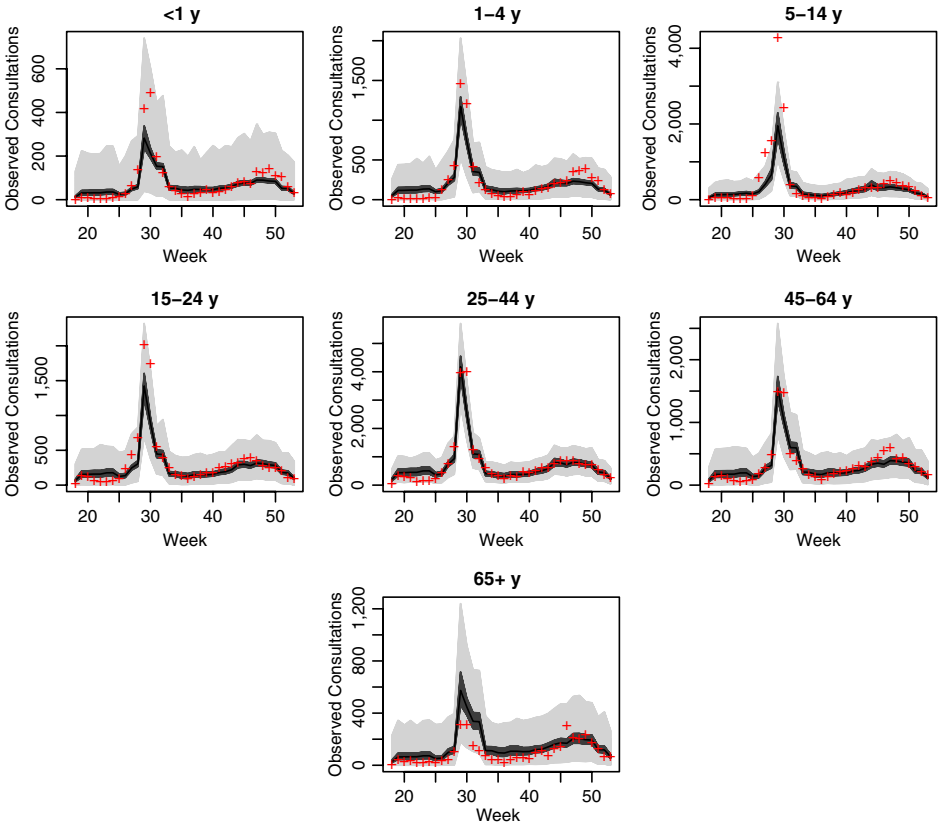

Fig. S2. Comparison of the observed GP consultation data (in red), against the expected number of consultations as predicted by the model (credible intervals shaded in black) and the attached negative-binomial error (shaded gray), which provide 95% credible intervals for the observed number of weekly consultations, by each age group in Greater London.

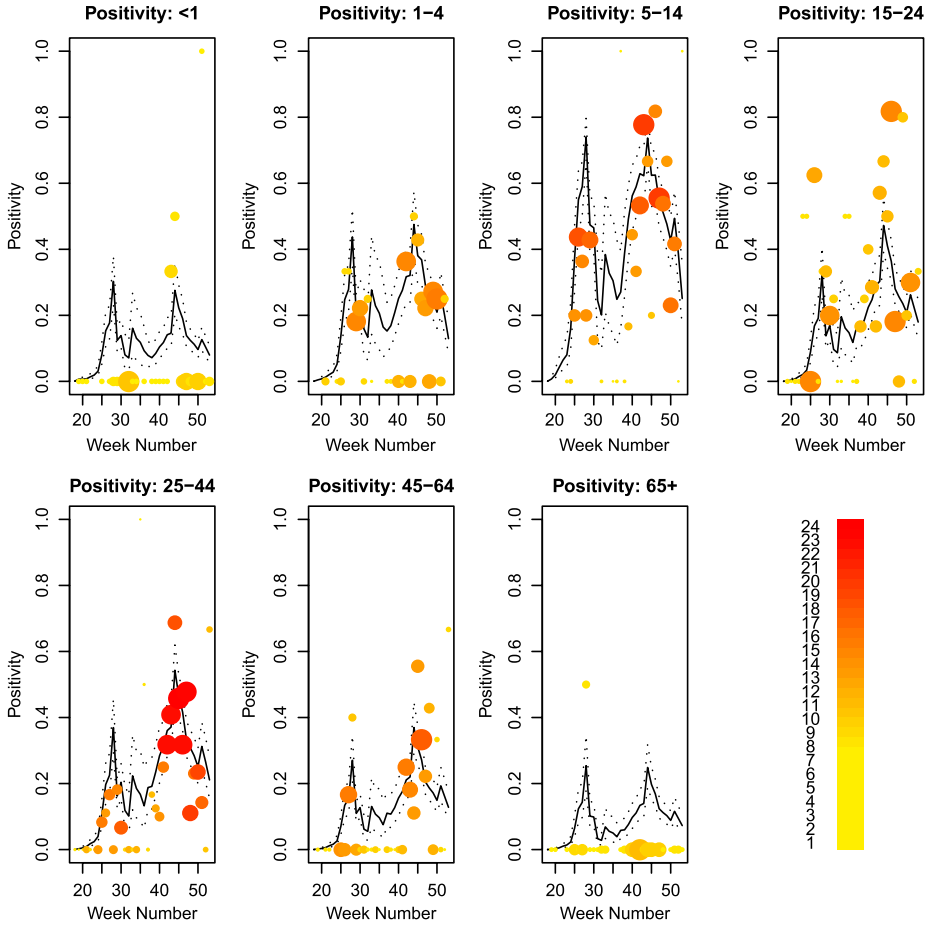

Fig. S3. Model fit to the virological positivity data. Dots represent observed virological positivity data by week, the size of the dot is proportional to the size of the virological samples relative to the other samples in the same plot, and the color an indicator of the absolute sample size. Again, each panel corresponds to the age group featured in the plot header.

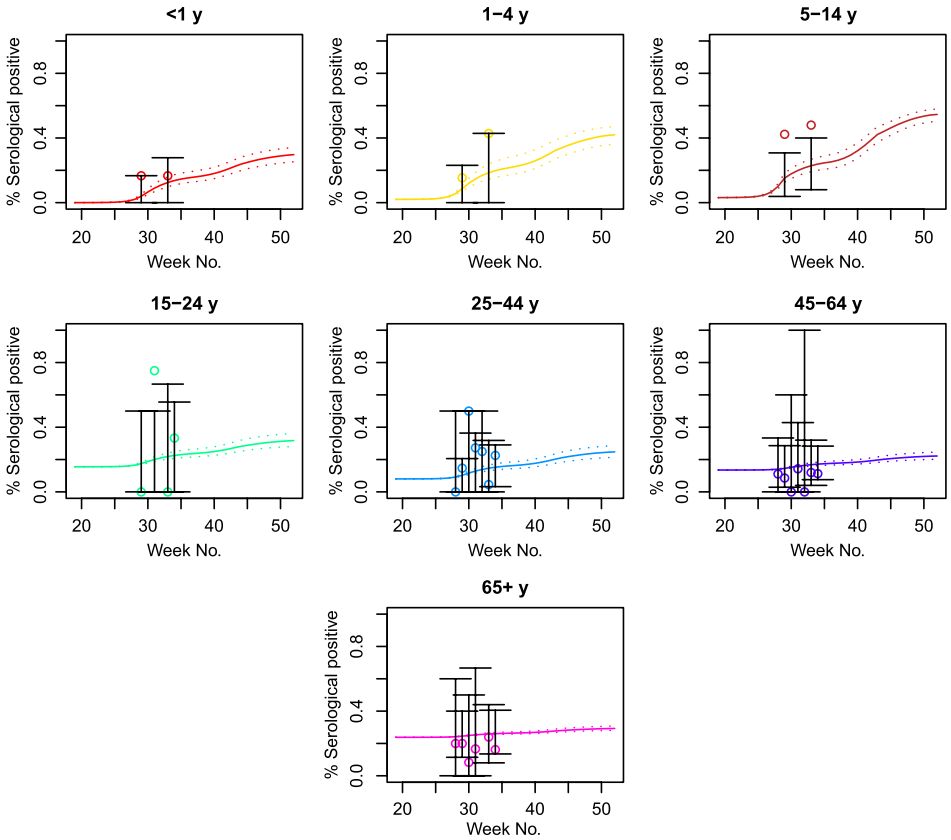

Fig. S4. Model fit to the serological positvity data, with data aggregated to weekly units. The points represent the observed weekly positivities (where samples exist); the curve represents the model-estimated positivity (together with a dashed 95% credible interval). The bars attached to each data point give credible intervals for the observed positivity under the model, conditional upon the sample denominators.

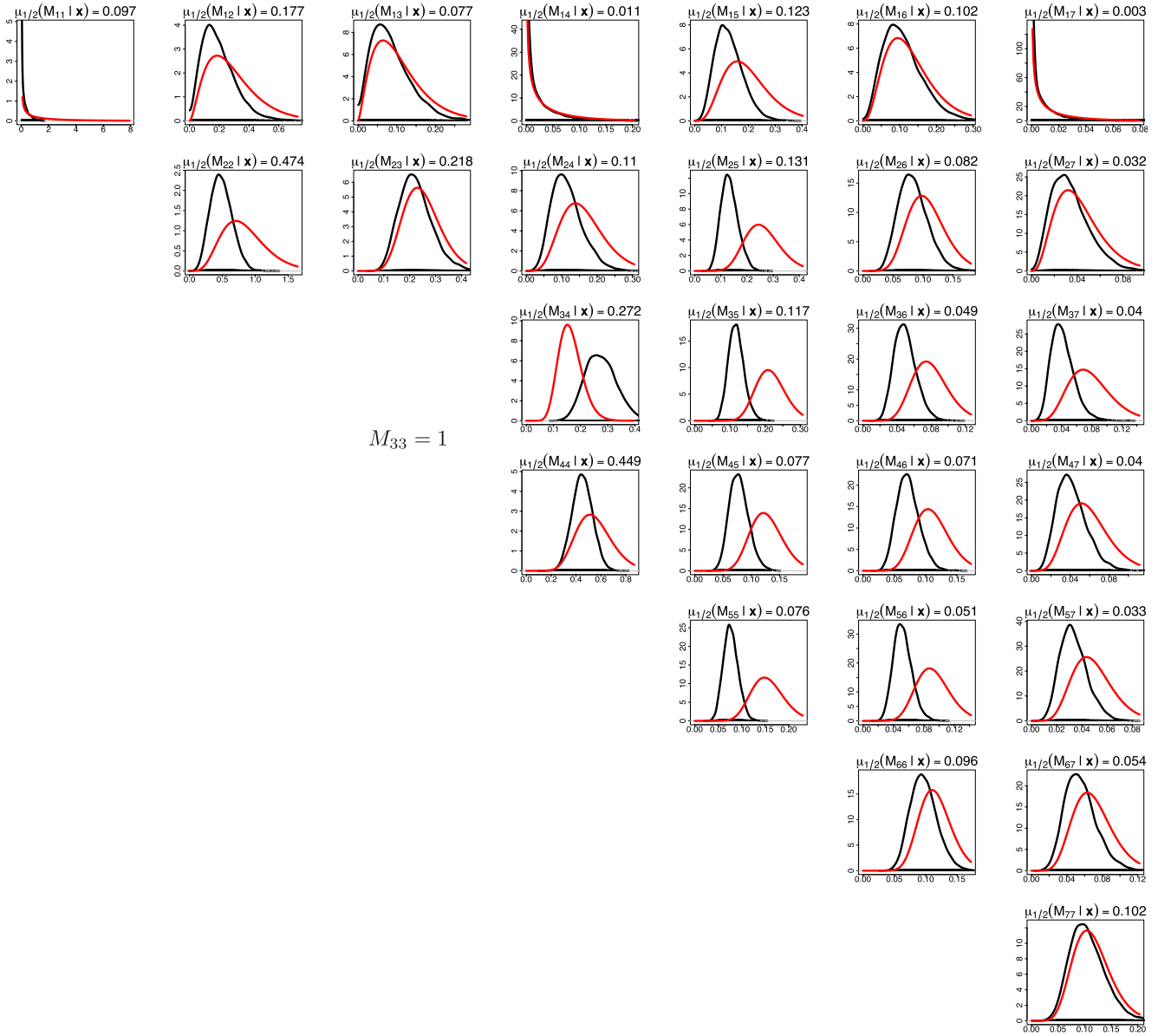

Fig. S5. Prior (in red) and posterior (in black) distributions for the parameters of the contact matrices. Posterior medians for the parameters are given in the headings of the individual plots.

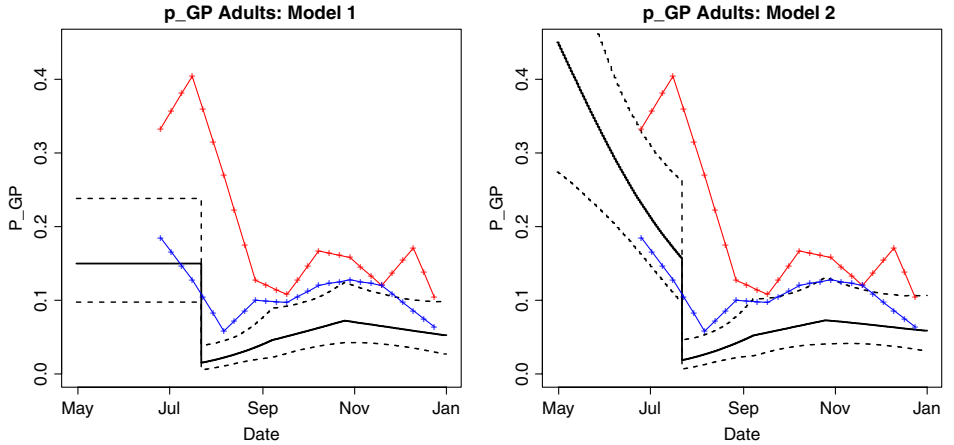

Fig. S6. Temporal evolution of the propensity to consult with a GP under the two piecewise logistic-linear models.

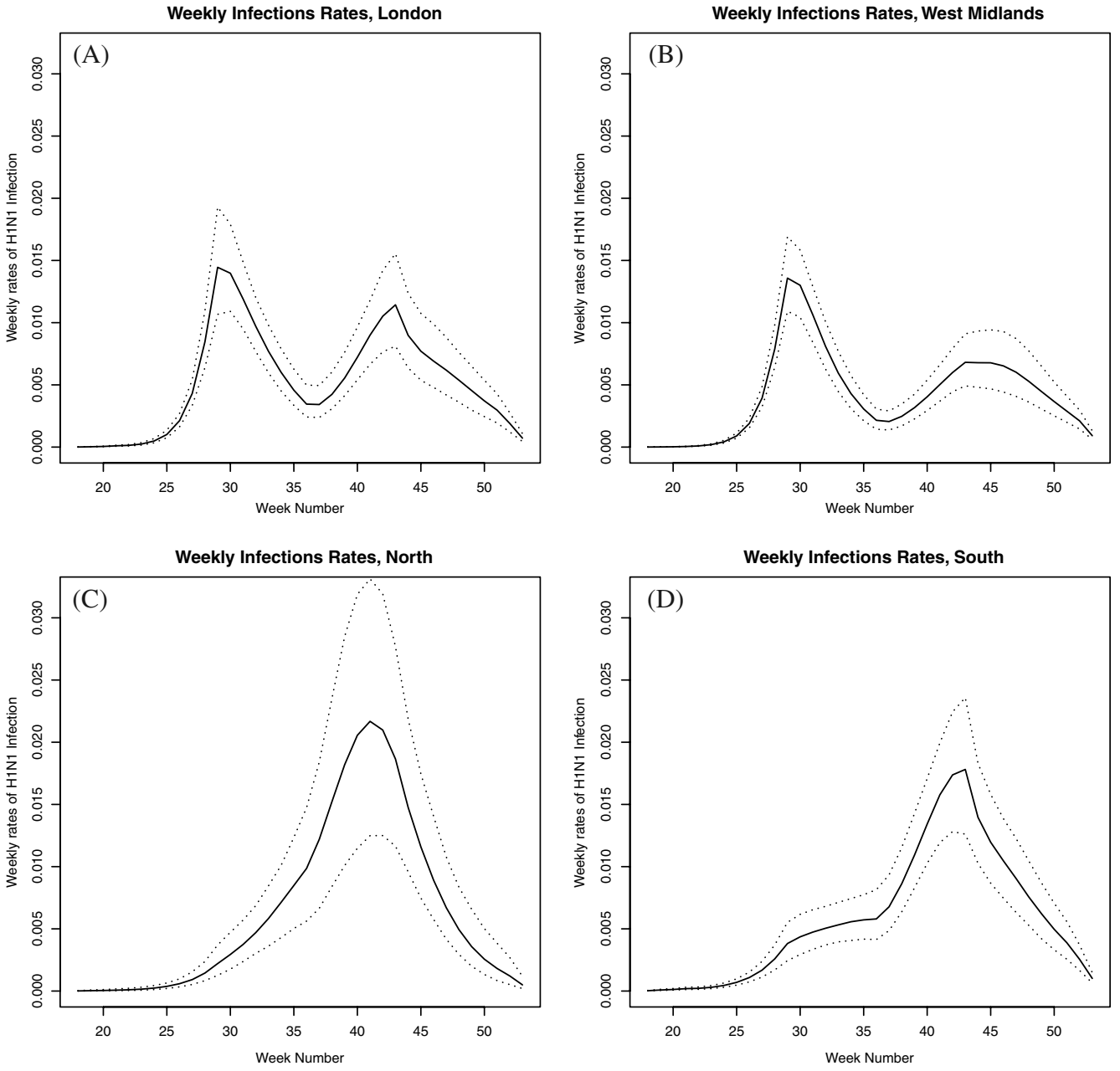

Fig. S7. Estimated epidemic curves and 95% pointwise credible intervals for the epidemic curves in the four regions: (*A*) London, (*B*) West Midlands, (*C*) North, and (*D*) South. For ease of comparison, incidence is expressed as a per capita rate.

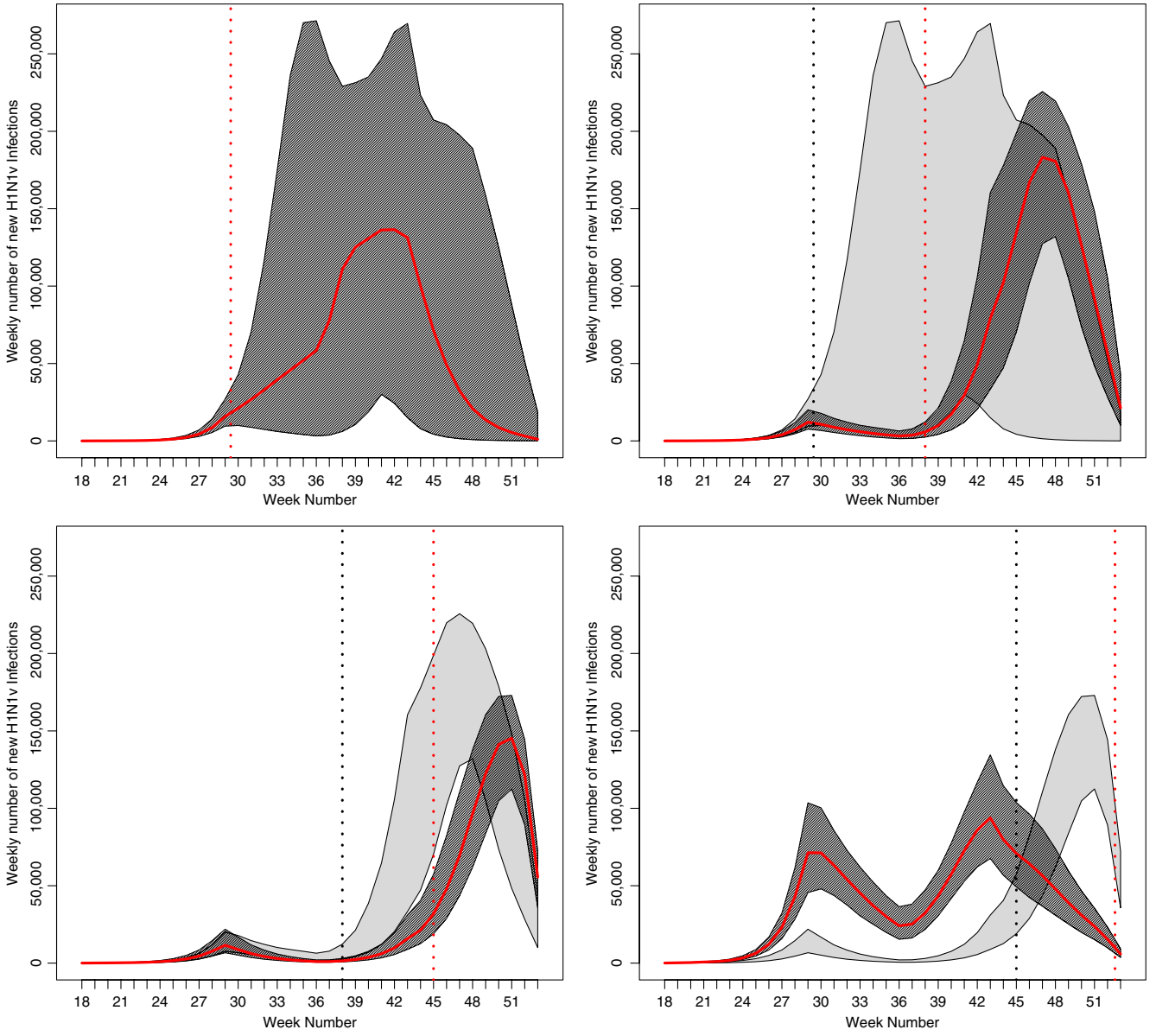

Fig. S8. Estimated epidemic curves and 95% pointwise credible intervals after 83, 143, 192, and 245 d of data excluding all serology.

Table S1. Assumed values and prior distributions for model parameters

Transmission model parameter	Prior/fixed value
Exponential growth rate, ψ_r	approximately $\Gamma(6.3,57)$
Initial log-hazard of GP consultation, ν	approximately $N(-19.15,16.44)$
Mean infectious period, d_I	$2+Z$, $Z\sim\Gamma(518,357)$
Mean latent period, d_L	2
Contact matrix parameters, m_i	approximately $U[0,1]\quad\forall\, i$
Initial proportion susceptible, π_a, in age group a	1(< 1 y), 0.980 (1–4), 0.969 (5–14), 0.845 (15–24), 0.920 (25–44), 0.865 (45–64), 0.762 (65+)
Disease and reporting model parameters	**Prior/fixed value**
Mean (SD) of gamma distributed incubation times	1.6 (1.8)
Proportion of infections who become symptomatic, θ	approximately $\beta(32.5,18.5)$
Proportion of cases who consult a GP, p_{GP}	$\log(p_i/1-p_i))\sim\begin{cases}N(-0.187,0.166) & i=1,2\\ N(0.426,0.929) & i=3,4\\ N(-0.319,0.263) & i=5,6\\ N(-0.284,0.264) & i=7,8\end{cases}$
Proportion of cases lab-confirmed, p_{CC}	approximately $\beta(1.03,2.69)$
Mean (SD) of gamma distributed waiting time from symptoms to GP consultation	2.0 (1.2)
Mean (SD) of gamma distributed waiting time symptoms to lab confirmation	6.6 (3.7)
Mean (SD) of gamma distributed reporting delay of GP consultations	0.5 (0.5)
Reporting delay of lab confirmations	0
GP consultation data dispersion parameters η_i	approximately $\Gamma(0.01,0.01)$, $i=1,2$

Table S2. Posterior medians and 95% credible intervals for some key epidemic quantities under different assumed values for the test sensitivity

Parameter	$k_{sens}=1.0$	$k_{sens}=0.9$	$k_{sens}=0.8$	$k_{sens}=0.5$
R_0	1.65 (1.56, 1.75)	1.64 (1.55, 1.73)	1.62 (1.54, 1.71)	1.56 (1.49, 1.64)
θ	0.328 (0.211 ,0.468)	0.354 (0.239, 0.503)	0.376 (0.253, 0.528)	0.441 (0.311, 0.594)
$\mathbb{P}_{\text{first wave}}$	0.885	0.892	0.924	0.976
$\text{IAR}_{\text{first wave}}$, %	8.90 (7.20, 10.9)	9.00 (7.20, 11.0)	9.00 (7.30, 10.9)	9.10 (7.50, 11.1)
$\text{SAR}_{\text{first wave}}$, %	2.90 (1.80, 4.30)	3.10 (2.00, 4.70)	3.30 (2.20, 4.90)	3.90 (2.80, 5.50)
$\text{IAR}_{\text{second wave}}$, %	10.4 (8.10, 13.0)	10.0 (7.70, 12.6)	9.80 (7.50, 12.4)	9.20 (7.00, 11.6)
$\text{SAR}_{\text{second wave}}$, %	3.40 (2.10, 5.10)	3.50 (2.20, 5.40)	3.60 (2.40, 5.60)	4.10 (2.80, 5.80)

$\mathbb{P}_{\text{first wave}}$ gives the probability that the peak of the first wave of infection is higher than the peak of the second wave. IAR gives the infection attack rate and SAR the symptomatic infection attack rate, with the subscripts indicating the wave of infection to which each attack rate refers.

Table S3. Posterior medians and 95% credible intervals for some key epidemic quantities under different contact patterns and propensities to consult with a GP

Parameter	Main analysis	M estimated	$p_{GP}(t,a)$ model 1	$p_{GP}(t,a)$ model 2
R_0	1.65 (1.56, 1.75)	1.58 (1.52, 1.66)	1.63 (1.55, 1.73)	1.66 (1.57, 1.75)
θ	0.328 (0.211, 0.468)	0.327 (0.215, 0.472)	0.322 (0.207, 0.459)	0.280 (0.183, 0.414)
$\mathbb{P}_{\text{first wave}}$	0.885	0.740	0.895	0.690
$\text{IAR}_{\text{first wave}}$, %	8.90 (7.20, 10.9)	9.60 (7.60, 11.6)	8.00 (6.20, 9.90)	7.60 (5.90, 9.40)
$\text{SAR}_{\text{first wave}}$, %	2.90 (1.80, 4.30)	3.10 (2.00, 4.80)	2.50 (1.60, 3.80)	2.10 (1.30, 3.10)
$\text{IAR}_{\text{second wave}}$, %	10.4 (8.10, 13.0)	11.7 (9.80, 13.8)	10.6 (8.20, 13.3)	11.7 (9.00, 14.4)
$\text{SAR}_{\text{second wave}}$, %	3.40 (2.10, 5.10)	3.90 (2.50, 5.70)	3.40 (2.10, 5.30)	3.30 (1.90, 5.10)

The table contains the analysis presented in the main text (column 1), an estimated contact matrix (column 2), piecewise-linear $p_{GP}(t,a)$. Rows have the same interpretation as in Table S2.

Table S4. Posterior medians and 95% credible intervals for some key epidemic quantities as estimated in the four disjoint regions

Parameter	London	West Midlands	North	South
R_0	1.65 (1.56, 1.75)	1.67 (1.61, 1.73)	1.43 (1.32, 1.54)	1.42 (1.35, 1.48)
θ	0.328 (0.211, 0.468)	0.416 (0.295, 0.560)	0.337 (0.214, 0.491)	0.327 (0.209, 0.477)
$\mathbb{P}_{\text{first wave}}$	0.885	0.999	0.00	0.00
m_1	0.469 (0.405, 0.539)	0.422 (0.372, 0.477)	0.949 (0.791, 0.998)	0.969 (0.870, 0.999)
m_2	0.535 (0.0395, 0.970)	0.590 (0.0408, 0.981)	0.603 (0.0391, 0.983)	0.460 (0.0207, 0.969)
m_3	0.279 (0.0323,0.481)	0.449 (0.297,0.550)	0.735 (0.337,0.964)	0.454 (0.138,0.713)
m_4	0.295 (0.0115, 0.919)	0.525 (0.0283, 0.977)	0.486 (0.0248, 0.972)	0.394 (0.0174, 0.963)
m_5	0.522 (0.276, 0.784)	0.940 (0.789, 0.998)	0.834 (0.369, 0.993)	0.284 (0.0215, 0.694)