


RESEARCH ARTICLE

Open Access



# Forecasting the 2017/2018 seasonal influenza epidemic in England using multiple dynamic transmission models: a case study

Paul J. Birrell<sup>1,2\*</sup> , Xu-Sheng Zhang<sup>2</sup>, Alice Corbella<sup>1</sup>, Edwin van Leeuwen<sup>2</sup>, Nikolaos Panagiotopoulos<sup>2</sup>, Katja Hoschler<sup>3</sup>, Alex J. Elliot<sup>4</sup>, Maryia McGee<sup>5,6</sup>, Simon de Lusignan<sup>5,6</sup>, Anne M. Presanis<sup>1</sup>, Marc Baguelin<sup>2,7,8</sup>, Maria Zambon<sup>2</sup>, André Charlett<sup>2</sup>, Richard G. Pebody<sup>2†</sup> and Daniela De Angelis<sup>1,2†</sup>

## Abstract

**Background:** Since the 2009 A/H1N1 pandemic, Public Health England have developed a suite of real-time statistical models utilising enhanced pandemic surveillance data to nowcast and forecast a future pandemic. Their ability to track seasonal influenza and predict heightened winter healthcare burden in the light of high activity in Australia in 2017 was untested.

**Methods:** Four transmission models were used in forecasting the 2017/2018 seasonal influenza epidemic in England: a stratified primary care model using daily, region-specific, counts and virological swab positivity of influenza-like illness consultations in general practice (GP); a strain-specific (SS) model using weekly, national GP ILI and virological data; an intensive care model (ICU) using reports of ICU influenza admissions; and a synthesis model that included all data sources. For the first 12 weeks of 2018, each model was applied to the latest data to provide estimates of epidemic parameters and short-term influenza forecasts. The added value of pre-season population susceptibility data was explored.

**Results:** The combined results provided valuable nowcasts of the state of the epidemic. Short-term predictions of burden on primary and secondary health services were initially highly variable before reaching consensus beyond the observed peaks in activity between weeks 3–4 of 2018. Estimates for  $R_0$  were consistent over time for three of the four models until week 12 of 2018, and there was consistency in the estimation of  $R_0$  across the SPC and SS models, and in the ICU attack rates estimated by the ICU and the synthesis model. Estimation and predictions varied according to the assumed levels of pre-season immunity.

**Conclusions:** This exercise successfully applied a range of pandemic models to seasonal influenza. Forecasting early in the season remains challenging but represents a crucially important activity to inform planning. Improved knowledge of pre-existing levels of immunity would be valuable.

**Keywords:** Transmission models, Seasonal influenza, Intensive care admissions, GP consultations, Nowcasting, Forecasting

\*Correspondence: [paul.birrell@phe.gov.uk](mailto:paul.birrell@phe.gov.uk)

†Richard G. Pebody and Daniela De Angelis contributed equally to this work.

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge Institute of Public Health, Robinson Way, Cambridge, CB2 0SR, UK

<sup>2</sup>National Infection Service, Public Health England, 61 Colindale Avenue, NW9 5EQ, London, UK

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

The evolution of influenza viruses results in annual epidemics of seasonal influenza, with less frequent global pandemics occurring due to the emergence of novel influenza viruses to which there is little population immunity. Both have the potential to cause substantial burden of disease to the population and to present significant challenges to already strained healthcare systems [1–3]. Many countries have invested substantial resources to developing epidemiological and virological surveillance tools to rapidly detect the onset of the influenza season each year [4, 5]; to measure the level of activity and the impact on the health service; and to characterise the main circulating virus strains and how well they match the seasonal vaccine. Attempts, however, to undertake seasonal influenza forecasting have been historically challenging due to the complex interplay between the influenza virus, population immunity, and environmental factors [6]. Nonetheless, progress has been made in the provision of short-term predictions of influenza in some countries in recent years particularly in the United States [5, 7], encouraging similar endeavours elsewhere.

In the UK, the National Risk Register of Civil Emergencies lists an outbreak of pandemic influenza as the greatest risk faced by its population [8]. Before, during and after 2009 A/H1N1pdm, quantitative approaches for real-time modelling and forecasting burden have been developed [9–11]. The availability of these models, together with complementary surveillance and data collection systems including sero-epidemiology for seasonal influenza, provided the opportunity to address the challenge of predicting seasonal influenza activity in England. This became a pressing need, when, following a particularly intense 2017 influenza season in Australia [12], prior to the winter season 2017/8 the National Health Service (NHS) put in place winter preparedness plans to manage potential acute pressures on the health service [13]. Ultimately, England experienced the most intense influenza season since the first post-pandemic season in 2010/11 [13]. Questions were raised about how the epidemic would evolve, when it would peak, how intense the peak in activity would be, and what would be the resulting demands on the health service, in terms of burden on GPs, hospitals, and intensive care units.

This study reports the advanced analytical and modelling experience at a national level led by Public Health England (PHE) during the 2017–2018 influenza season and in particular the attempt to undertake short- and medium-term forecasts of influenza activity and the impact of influenza on the health care service.

## Methods

### Data

The annual monitoring of influenza activity in England is based on a series of data streams. Those specifi-

cally utilised for the purposes of this work were weekly and daily consultations for influenza-like illness (ILI) in general practice (GP); virological testing of respiratory swabs obtained from patients consulting in GP; laboratory confirmed influenza admissions to hospitals and intensive care and high dependence units (ICU/HDU); and population-based strain- and age-specific serological data on influenza immunity. These data sources have been described in detail elsewhere and we review them only briefly here [14–17].

The GP ILI data represents patients attending primary care with acute ILI, a proportion of which will be due to influenza infection. Data were obtained from two sources: the Weekly Returns Service of the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), a sentinel GP network covering a weekly population (in December 2017) of approximately 2 million people from over 200 practices [18]; and the PHE GP in-hours-syndromic-surveillance system, which collates daily ILI consultations stratified by, amongst other things, age group and NHS region from approximately 3,500 GP practices, representing about 50% of the total practices in England [16]. Both of these data sources are routinely available throughout the year, with a period of enhanced influenza surveillance starting at week 40, and ending at week 20, irrespective of the amount of circulating influenza. The RCGP RSC sentinel network has integrated virological monitoring [18], whereby GPs undertake respiratory swabbing of a subset of patients consulting for ILI. Swabs are tested by reverse transcription-polymerase chain reaction to identify the component of ILI due to each influenza strain through the presence of type and subtype specific positive swabs in the sample.

Data on influenza confirmed hospitalizations and intensive care admissions are collected through the UK Severe Influenza Surveillance System (USISS) [15, 19]. Weekly numbers of laboratory-confirmed influenza cases (of all the commonly circulating strains: A/H1N1pdm09, A/H3N2, B) admitted to an ICU or high dependence unit (HDU) and the number of confirmed influenza deaths in ICU/HDU are reported from all NHS trusts in England from week 40 to week 20 of the following year, alongside attendant information including age and influenza subtype. In addition to this mandatory scheme, a subgroup of NHS trusts in England is recruited every year to participate in the USISS sentinel scheme [20], which provides weekly numbers of laboratory confirmed influenza hospitalisations (A/H1, A/H3 and B).

A final source of information is provided by intra-seasonal cross-sectional population-based serological survey data. These data provide a measure of the seroprevalence of strain-specific antibodies to A/H3, A/H1 and B influenza viruses using haemagglutination inhibition (HAI) assays with an HAI titre > 1/40. These data inform

the susceptibility of the population after the end of the previous season, but prior to both the seasonal influenza vaccination campaign (which should boost population immunity) and the onset of seasonal virus circulation.

### Influenza transmission models

Four deterministic compartmental transmission models were employed to estimate and forecast the evolution of the 2017/18 seasonal influenza in England: a Stratified Primary Care Surveillance Model (SPC); a Strain-Specific model (SS); a Severity-based model (ICU); and a Synthesis model. The common transmission structure is of the Susceptible (S), Infectious (I), Removed (R) type, adapted for greater realism to include additional Exposed (E) or I compartments in some cases. All the models assume homogeneous mixing between susceptible and infected individuals within population strata. Transmission dynamics are linked to the observed data through appropriate disease reporting models. The data streams used to estimate model parameters and relevant epidemic quantities vary according to the model. Detailed information on the models' structure, data used, distributional assumptions and estimation approaches is given in the [Web Extra Material](#) (Sections A–B; Table S1) while here we give a short summary of each.

**Stratified Primary Care Model** This is a SEIIR model, a modified version of the model developed to reconstruct the 2009 H1N1 influenza pandemic [11, 21, 22]. The model uses daily (or weekly) data on the number of GP consultations for ILI from the PHE influenza surveillance dataset augmented with the RCGP's virological data to estimate the component of ILI due to influenza. In the first week of analysis, the number of ILI consultations due to influenza were simply estimated by multiplying the total number of consultations in sentinel RCGP practices by the proportion of swabs testing positive for an influenza virus, obtaining what we term ILI+. When more detailed data became available, it was possible to model jointly daily ILI counts and swab positivity data, appropriately accounting for the size of the virological samples, and to specify distinct epidemic models in each of five regions in England [22]. A definition of the regions used is available in the [Web Extra Material](#) (Section B.1.1). When using ILI rather than ILI+, it is necessary to account for "background" rates of consultation, the component of ILI not attributable to influenza. This level of consultation is estimated by fitting an endemic/epidemic model [23] to 3 years of historic ILI data prior to the 2017/2018 flu season.

**Strain-Specific Model** The SS model has an SIR structure for each of the three influenza strains (A/H1, A/H3 and B), which are assumed to transmit independently within the population [24]. The model uses weekly GP

consultations for ILI from the RCGP RSC, together with the strain-specific virological data from the RCGP RSC's virological monitoring to identify the contribution of each strain to the overall number of consultations. Serological data were used to inform strain-specific susceptibility at the start of the season.

**ICU-based Model** The ICU model for seasonal influenza represents a development of an existing influenza SEIIR transmission model, which uses intensive care influenza admissions from the USISS system [25]. The transmission rate is allowed to vary over time to account for school holiday periods and the possible effects of changing patterns of interaction between age groups [26]. Transmission dynamics are linked to the observed ICU admissions through a delay from infection to admission and through the assumption that only a proportion of the infections is admitted to ICU.

**Synthesis Model** The synthesis model uses three data streams to estimate the underlying transmission dynamics: ICU and hospital admissions from USISS and the ILI+ dataset derived from RCGP surveillance. Transmission dynamics are described by a SEIIR model assuming a constant transmission rate and random mixing. The basic feature of the disease reporting process are taken from previous studies [15, 21, 25].

### Estimation of parameters and quantities of interest

Using weeks as defined by the International Organisation for Standardization (ISO), the above models were run weekly in each of weeks 1 to 12 of 2018, with all analyses covering the period from week 40 of 2017 (denoted 2017w40), which started on the 2<sup>nd</sup> October. For each week and each model, we estimate model parameters (see Table S1 in the [Web Extra Material](#)) in a Bayesian framework. Quantities that can be estimated from all models and have a common interpretation were the focus of comparisons. In particular, we estimated features of the epidemic that are of great public health interest, including the timing and magnitude of the peak in the burden on healthcare services (influenza-related visits to GPs, hospital and ICU admissions);  $R_e(t)$ , the effective reproduction number, representing the average number of infections generated by a single infection in the population (see [Web Extra Material](#)); and the propensity of infected cases to interact with the particular healthcare services (GP, Hospital, ICU). In deterministic models, the reproduction number decays over time with the depletion of susceptibles, so estimates of its value at the start of the influenza season only are presented and this quantity is denoted  $R_e \equiv R_e(0)$ .

## Forecasting

From the estimates above it is possible to derive estimates of burden in terms of new infections, GP consultations and admissions to secondary care over the coming weeks. However, these estimates refer to quantities that are not directly observable from available data. Even GP consultations and data on admissions for severe disease typically refer only to a sub-sample of the whole population, representing a noisy version of these true underlying quantities. To assess the ability of the models to make predictions, it is necessary to contrast the model-based forecasts for observable quantities with the corresponding subsequent observations. We carried out such assessment by constructing posterior predictive distributions for future data points and plotting them alongside the observed values [7, 27].

## Results

### Parameter estimation

Detailed information on estimates from all models including uncertainty can be found in Table S2 in the [Web Extra Material](#). Here we compare estimates (and 95% Credible Intervals) of common parameters from the different modelling approaches as they are estimated over successive weeks. Note that as not all models provide comparable outputs, each panel of Fig. 1 refers only to the relevant models. Panel A reports estimates of the overall effective reproductive number  $R_e$  from three of the models (SPC, ICU and Synthesis), whereas Panel B refers to the strain-specific  $R_e$  obtained from the SS model. The SPC model persistently estimates the highest  $R_e$  in Panel A, with the central estimate comparable to the strain-specific estimates in Panel B. Estimates from the Synthesis and ICU models seem also to be consistent, converging to very similar values (1.25–1.26) by ISO week 12 (Panel C). Only the SS model made use of the population intra-seasonal seroprevalence data.

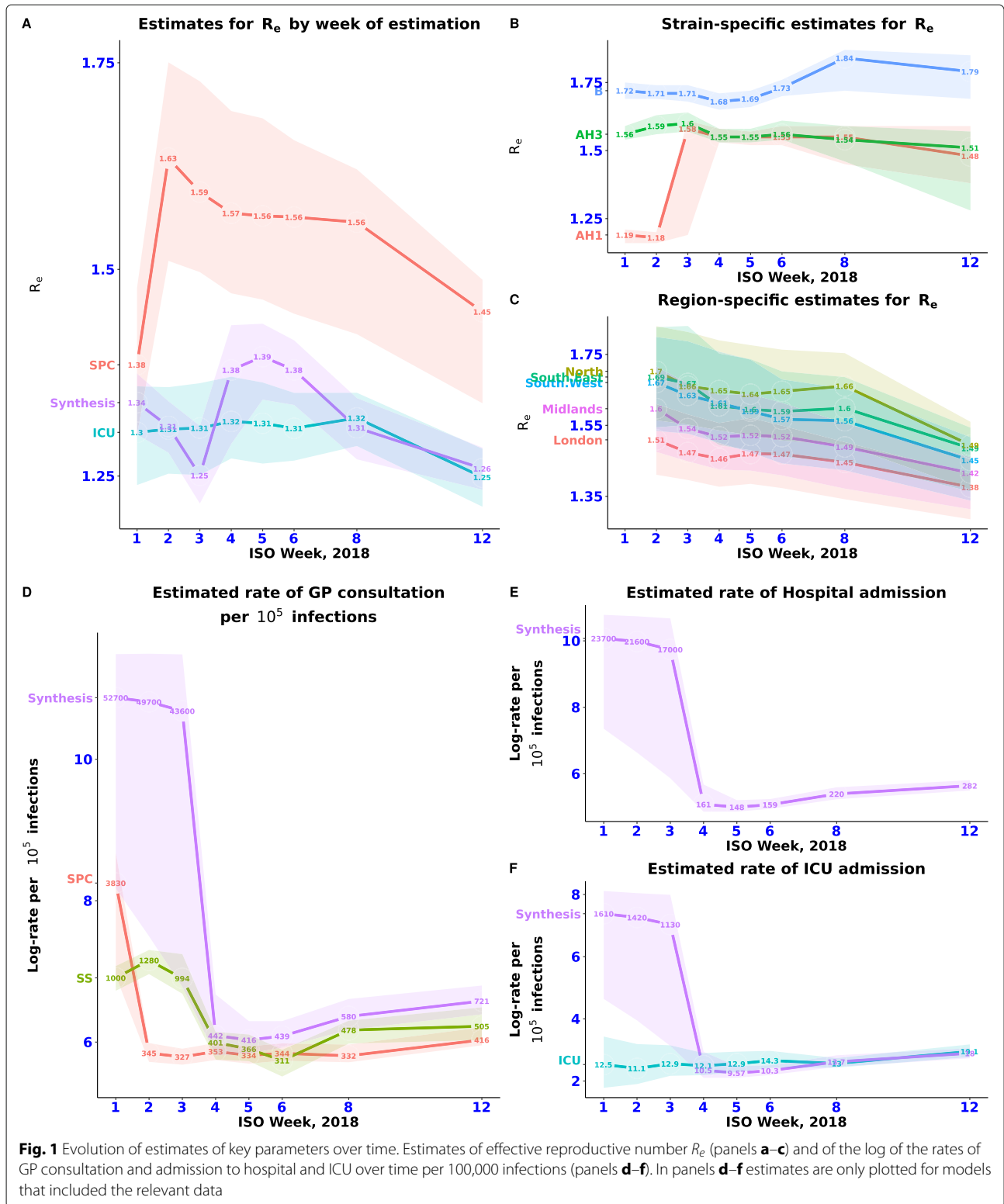
Panels D–F report estimates of health care seeking rates per 100,000 infections, of GP consultation for ILI (Panel D), hospitalisation (Panel E) and ICU admission (Panel F). These rates are not estimated if a model doesn't utilise the relevant dataset. In Panel D, after an initial volatile pattern, the estimates settle down around 2018w4. In the SPC model, the ILI+ data used in 2018w1 are replaced from 2018w2 by separate GP ILI and swab positivity datasets. As a result, the estimated rate of GP consultation drops sharply and becomes much more precise. For the SS and SPC models, estimates of the propensity to consult in GP stabilise, over time, at around 0.5%. Results from the Synthesis model of all three healthcare seeking events, suggest a slightly higher value for the GP consulting rate, a figure of around 0.3% of infections leading to hospitalisation, with a lower estimate, below 0.02% of infections requiring ICU admission, in agreement with the

ICU model whose estimates are the most consistent over the period.

### Estimating timing and level of peak activity in primary and secondary care

Figure 2 displays estimates (posterior means), with uncertainty (95% Credible Intervals) of the timing of the peak and the peak intensity for the influenza-attributable ILI consultation rates and ICU admission rates by model and analysis time. The grey-shaded band on each plot gives the period of time over which peaks in the datasets were observed. Specifically, Panel A shows the evolution of the estimate of the peak week in GP ILI consultations as data accumulate, from some initial, quite heterogeneous, estimates, to a consensus estimate. For the GP consultation rates this consensus, placing the peak between 2018w3 and 2018w8, is only reached after 2018w4. Before that the uncertainty attached to the initial estimates varies substantially by model: the Synthesis model provides volatile results, while both the SPC and SS models are giving by this time a more precise indication on the position of the peak. The SPC model's estimates seem overly precise, perhaps due to a lack of uncertainty in the estimation of the non-influenza ILI consultation rates (see Section B.1 in the [Web Extra Material](#)). Note also that the SPC model is able to provide region-specific estimates (Panel B). A similar message comes from the estimation of peak in the ICU rates, with an initial uncertainty diminishing after 2018w4, once a peak in the ICU data has definitely occurred.

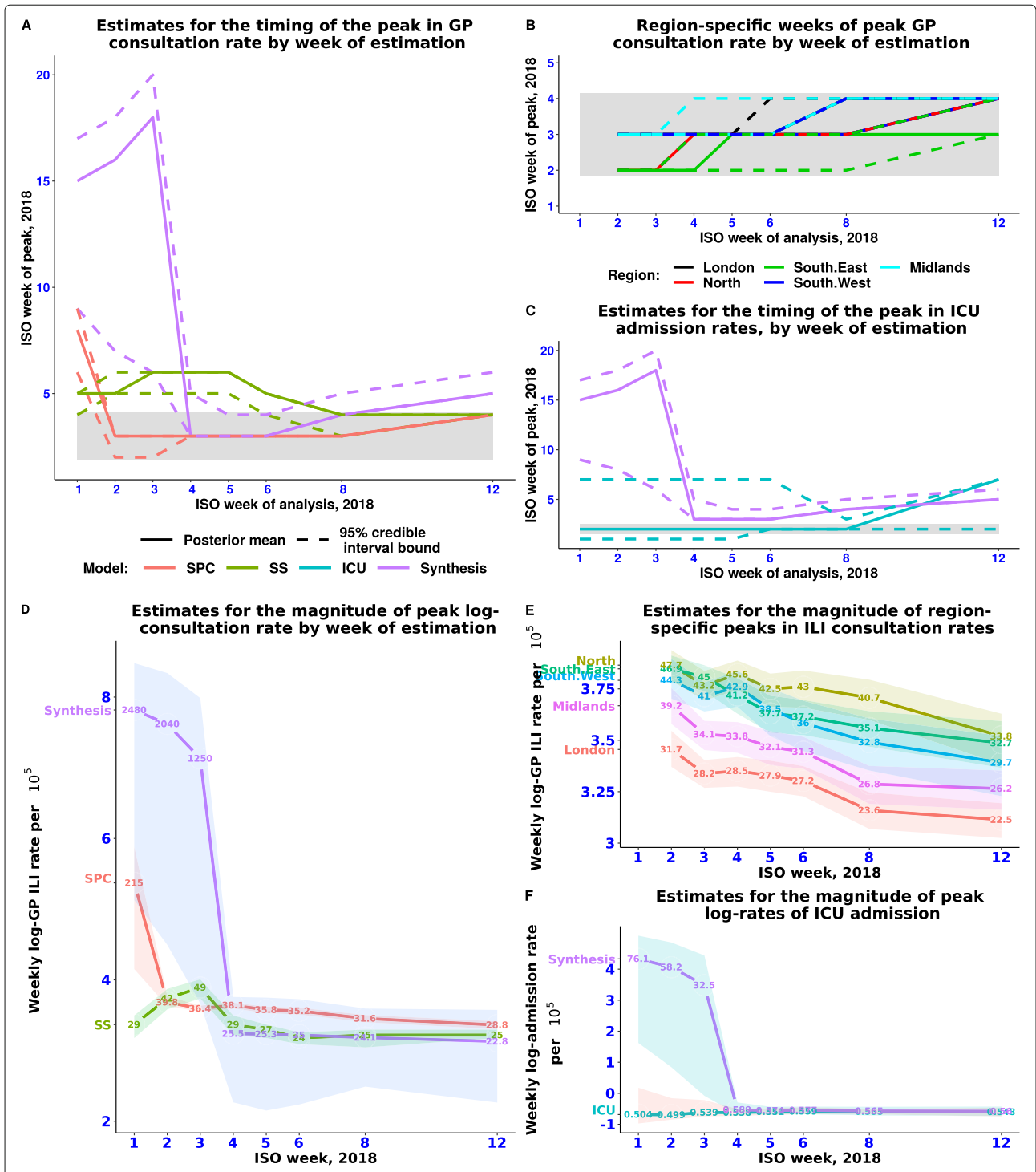
In estimating the magnitude of the peak in GP consultations (Panel D), both the Synthesis model and the SPC model give initially a large estimate (based on ILI+ data). The SPC model prediction drops quickly in 2018w2, whereas the Synthesis model takes two additional weeks to do so. Over time the magnitude of the peak GP activity estimated under each of the three models gradually declines, with the SPC model continuing to estimate slightly higher peak activity. Although more consistent over time, the estimates for the peak intensity in the SS and SPC models again seem to be very precise given the predicted fall in activity over time. In Panel E, the regional estimates for the peak intensity behave in much the same way as the national estimates for  $R_e$  with a gradual decline over time, with London and the Midlands having the lowest rates of peak activity and the North and South-East the highest. Estimates for the peak levels of ICU activity (Panel F) from the ICU model are very consistent over time, with the estimates from the Synthesis model converging to similar values from 2018w4 onwards. Table 1 presents the forecasts for the attack rates, the cumulative incidence of GP consultation, hospital and ICU admission over the course of the influenza season. These forecasts behave similarly to the forecasts for peak incidence (see Fig. 2d–f), in that they take a similar amount of time to



converge to stable values. However, over time, the attack rate forecasts by the ICU model are steadily increasing (due to a slower than anticipated decline in the number of admissions), whereas those for the SS model are

decreasing. There is little overlap between credible intervals from week 5 to week 12 across all models except the Synthesis model where forecasts are significantly less certain. Forecast ILLI+ attack rates from the SS model tend to





**Fig. 2** Evolution of forecasts of peak activity over time. Estimates for the peak timing (Panels a–c) and magnitude (Panels d–f) in GP consultations and ICU admissions over time. In Panels a–c the shaded grey region gives the time interval over which the observed peaks in ILI (and ILI+, Panels a and b) and ICU admission (Panel c) occur. In places, the credible intervals vanish, where the peak week is predicted with certainty to be in a specific week. Again, region-specific estimates are presented in panels b and e, while only the ICU and Synthesis models forecast the ICU admission rates and only these models feature in Panels c and f

be higher than those from the SPC and Synthesis models, whereas forecasts from attack rates of ICU admission are comparable between the ICU and Synthesis models.

The Synthesis model does, however, appear to estimate a high ratio between the number of hospital admissions and GP consultations, with there being less than three

**Table 1** Forecast attack rates

ISO Week	SPC model	SS model	ICU model	Synthesis model		
	ILI+	ILI+	ICU	ILI+	Hosp.	ICU
1	1650	365	4.54	19400	8730	594
	(456, 2660)	(315, 424)	(2.36, 10.8)	(163, 40700)	(738, 19200)	(50.1, 1320)
2	188	485	4.10	18500	8020	528
	(175, 206)	(415, 549)	(2.73, 8.13)	(1580, 33700)	(710, 15200)	(46.6, 1010)
3	174	570	4.74	8460	3320	220
	(165, 183)	(473, 648)	(3.57, 7.80)	(134, 20600)	(74.1, 8270)	(4.88, 555)
4	185	370	4.53	197	71.7	4.63
	(176, 194)	(322, 417)	(3.76, 6.04)	(61.1, 441)	(55.8, 128)	(3.54, 8.49)
5	175	348	4.80	190	67.4	4.35
	(167, 184)	(306, 378)	(4.11, 6.15)	(68.0, 331)	(57.2, 82.1)	(3.64, 5.42)
6	180	300	5.29	195	70.6	4.56
	(171, 189)	(262, 345)	(4.63, 6.40)	(71.4, 329)	(61.0, 83.1)	(3.89, 5.47)
8	173	280	5.01	225	85.3	5.29
	(164, 182)	(236, 305)	(4.68, 5.44)	(110, 354)	(70.9, 104)	(4.38, 6.53)
12	194	265	6.16	250	96.3	6.13
	(185, 202)	(249, 302)	(5.50, 6.97)	(108, 400)	(81.0, 115)	(5.10, 7.42)

Posterior median (with 95% credible intervals underneath) for the ILI+ attack rate (per 100,000 people), Hospitalization attack rate (per 100,000 people), and ICU attack rate (per 100,000 people). These attack rates are forecasts for the cumulative total number of GP consultations/Hospitalisation/ICU admissions attributable to influenza over the course of the whole influenza season. The 'ISO Week' gives the week of 2018 in which the forecast is made

GP consultations per Hospitalisation, suggesting some possible unaccounted-for bias in the data sources.

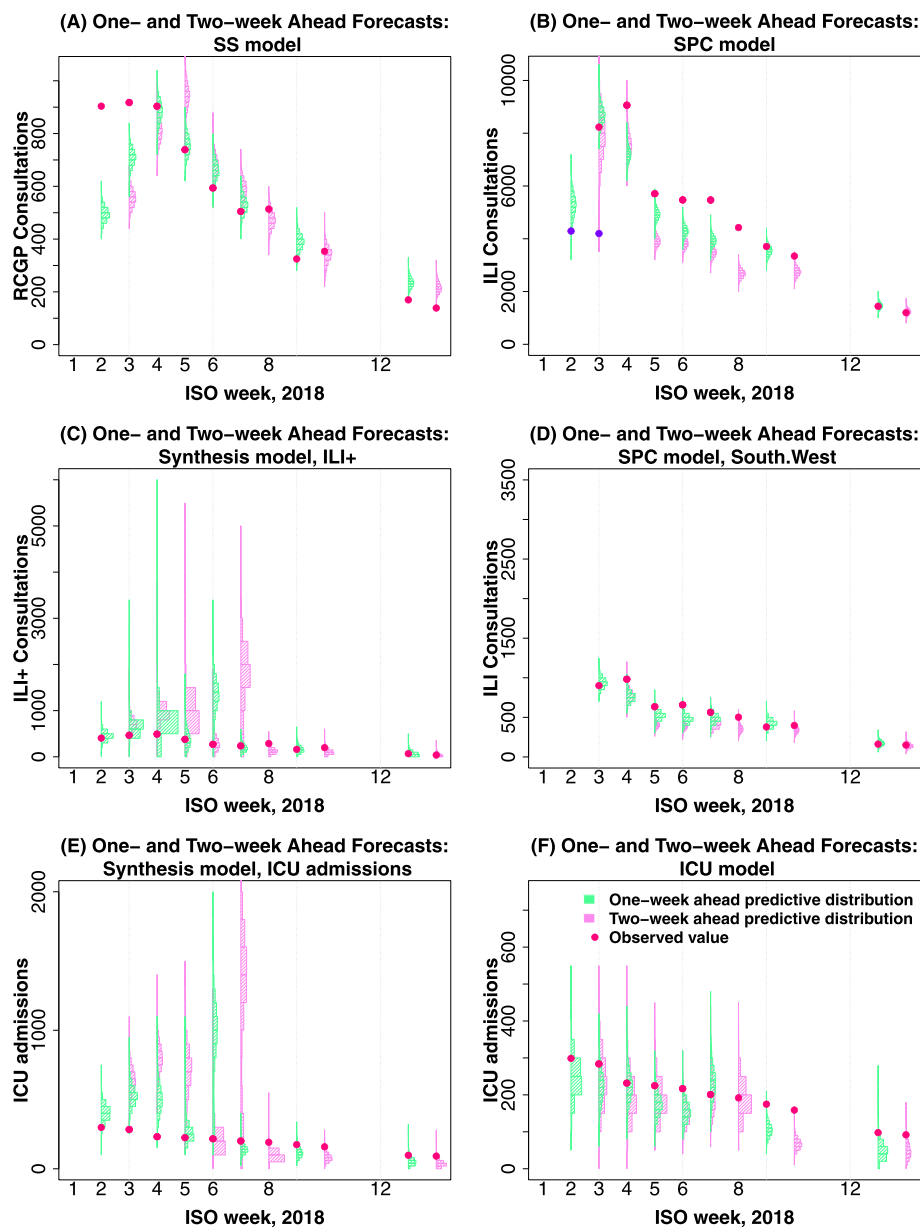
#### Forecasting activity and impact including timing of the peak

Moving from estimation to forecasting, Fig. 3 illustrates the ability of each model to forecast relevant quantities. Here, at each time in each plot, one-week ahead probabilistic forecasts made in the previous week (in green) and two-week ahead forecasts made two weeks prior (in pink) are plotted alongside the data point subsequently observed at that time (red dots, with blue dots in Panel B corresponding to the ILI+ data that the SPC model used in week 1). Panel A displays the GP ILI forecasts from the SS model, showing how the model struggles to anticipate the peak in observed consultations early in 2018 and how the forecasts improve over 2018w5–9, after peaks in activity have been observed. Forecasts from the SPC model (Panel B) show a different pattern, displaying a better forecasting ability in the early period, which degrades over 2018w5–9, before very accurately predicting the GP ILI activity in 2018w12. This lack of forecasting ability over 2018w5–9, however, is due to the overly precise estimation observed above. Across each of the individual regions, the observed ILI consultations over 2018w5–9 fall in the upper tail of the posterior predictive distributions (as exemplified in Panel D for the South West), but they combine to give a national total ILI consultation that lies comfortably above

the predictive distribution as seen in Panel B. The Synthesis model forecasts for the ILI data are shown in Panel C. The one week-ahead forecasts appear to perform reasonably well, but there is often high uncertainty attached to the two-week ahead forecasts, at times at odds with the one-week ahead forecast produced the following week. A similar big uncertainty is observed when using the Synthesis model to forecast ICU admissions (Panel E). Finally, the ICU model seems to exhibit good one-week and two-week forecasting performance, with almost all the data points lying within appropriate predictive intervals. However, it is to be noted that none of the observed values lie in the lower tail of the predictive distributions, suggesting some likely under-estimation of the ICU admissions over time.

#### Discussion

This study shows how models designed for pandemic influenza could be adapted to answer questions on unfolding seasonal flu activity, in particular when it will peak, what the level of peak activity will be and what is the health service impact in terms of hospital and ICU admissions. We addressed these questions through a number of available transmission models, each sharing a common population compartment structure, that used a range of different standard data streams to make sequential estimation of disease spread, of case/severity indicators, and of future epidemic activity, all as data accumulated over time. Estimates were typically consistent across models



**Fig. 3** One- and Two-week ahead forecasting performance of the models Credible intervals for one week ahead (green) and two week ahead (pink) forecasts for the data used by each model. The red dots represent observed data, with blue dots in Panel **b** giving observed ILI+ observations (which were not used for model fitting beyond 2018w1)

using similar information (e.g. the SPC and SS models) and generally more stable towards the latter stages of the season. There is a high-degree of volatility in estimates that rely on the ILI+ data (Synthesis model and SPC model in week 1). This is most likely due to the ILI+ data being reliant on virological swab positivity data that only has a relatively small number of positives. These data then represent a highly noisy signal of flu activity that could easily conflict with the other datasets used by the Synthesis model. Estimation of timing and magnitude of peak

activity was particularly challenging, as heterogeneous estimates were obtained from the different models in the early stages prior to observing peaks in the data.

To understand the predictive ability of the models, one- and two-week ahead forecasts were produced from each model and contrasted with the corresponding quantity subsequently observed. In this regard, models using a single data stream were typically more reliable, with prediction performance improving over time. More formal methods for assessing the forecast ability of a model exist,



which properly account for both coverage and precision of the forecasts [27]. While interesting to investigate, these are more appropriate to situations when comparing different models forecasting the same indicators, unlike here. Informally assessed, the SPC model which incorporated daily surveillance data was best able to forecast peak activity, highlighting the importance of daily data streams for nowcasting and forecasting purposes.

Each of the models used have some shared characteristics and consequently some shared limitations. They each have at their heart deterministic transmission structures relying on assumptions of heterogeneous mixing (in the SPC model this heterogeneity is within regions, not England-wide). Deterministic dynamics are suitable for capturing pandemic dynamics when there is a single circulating influenza strain to which the population is almost entirely susceptible. With seasonal influenza both environmental and demographic stochastic effects will be influential, particularly in absorbing any lack of fit. Here, for example, the single-strain models struggled to adapt to the patterns in the data that are not characteristic of SEIR-type dynamics and in particular the slow decline of influenza activity following the peak. This led to increased estimates of attack rates in 2018w12. Conversely, the deterministic dynamics of the strain-specific model couldn't account for the presence of low-levels of influenza A/H1 activity without eventually leading to more widespread transmission. The increased flexibility in stochastic transmission models could potentially absorb this lack of fit without biasing estimates to the same extent.

A sensitivity analysis was conducted to understand the role that serological data can have in the monitoring of seasonal influenza (see [Web Extra Material](#)). The initial susceptibility of a population, which is a parameter in the models used here, cannot be estimated through the use of the surveillance streams alone. External knowledge on levels of population immunity prior to the start of the season needs to be provided, knowledge to which estimates of infection spread, attack rates, and case/severity indicators are found to be highly sensitive. The sensitivity study showed that while  $R_e$  estimates are robust to the levels of initial susceptibility, estimates for the transmissibility (measured by  $R_0$ , see [Web Extra Material](#)) are inversely proportional to the susceptibility. Assuming higher initial susceptibility, the unaffected estimates of  $R_e$  will lead to higher rates of infection, as has been shown elsewhere [28]. To be able to explain the observed data there is a consequent impact on the case/severity ratios, which are estimated to be lower, leading here to differing estimates of (earlier) timing and magnitude of (higher) peak activity if early in the season. To weaken our strong assumption of homogeneous mixing, serological data can be of even further value if they are sufficiently

representative (by age, strain, region etc.) to permit corresponding stratification of the transmission model. Additionally, the contribution of influenza vaccine programmes (in terms of uptake and effectiveness) to susceptibility also needs to be taken into account, as the vaccine campaign occurs after the intra-seasonal sero-survey is completed and will clearly further affect the population sero-profile prior to the start of the season. This will have been of particular importance in 2017/18, due to the apparent impact of vaccine-related egg-adaptation, which may have reduced vaccine effectiveness against the circulating influenza A(H3N2) strains [29]. Further work on the potential role of sero-epidemiology to improving forecasting needs to be explored [30, 31] and a pilot study is ongoing [32].

## Conclusions

This exercise constituted a first attempt at establishing routine short-term forecasting of seasonal influenza activity in the UK with the aim of informing health service planning during the winter. The UK has a strong, integrated influenza surveillance system which provides an important opportunity to develop such approaches. Although modelling in the earlier stage of the season remains particularly challenging this has been a very valuable enterprise, identifying the key information requirements and the optimal modelling approaches. Reliable short-term predictions, particularly at local levels, for the number of cases in primary and secondary care can enable health service planners to optimally deploy limited capacity (e.g. hospital bed management). The work described here has identified the further developments required to achieve this: the use of more detailed serological data, the incorporation of information on vaccine coverage and building in additional flexibility to models to give less precise forecasts due to the presence of possible biases.

## Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12889-020-8455-9>.

**Additional file 1:** Web extra material. Appendix to this manuscript containing technical details of each of the models discussed here, some additional results and a sensitivity study on the impact of serological information.

## Abbreviations

GP: General practice; HAI: Haemagglutination inhibition; HDU: High-dependence unit; ICU: Intensive care unit; ILI: Influenza-like illness; ILI+: Influenza-like illness multiplied by virological swab positivity; ISO: International organization for standardization; NHS: UK national health service; PHE: Public health England; RCGP: Royal College of general practitioners;  $R_e$ : Effective reproduction number; RSC: Research and surveillance centre (of the RCGP); SEIR: Susceptible-exposed-infectious-recovered (with variation in the number of 'E' and 'I's used); SPC: Stratified primary care surveillance model; SS: Strain-specific; USISS: UK Severe Influenza surveillance system

### Acknowledgements

We acknowledge support from TPP and participating SystmOne practices and University of Nottingham, ClinRisk, EMIS and EMIS practices submitting data to the QSurveillance database and the PHE Real-time Syndromic Surveillance Team (GP in-hours). We also thank patients of the RCGP RSC practices who consented to having flu swabs taken and the RSC practices themselves for processing and sharing their data. The views expressed are those of the author(s) and not necessarily those of the MRC, PHE, or the NIHR.

### Authors' contributions

PJB, XSZ, RGP, DDA contributed writing for the manuscript; PJB, XSZ, AC, EvL prepared figures and tables; PJB, XSZ, EvL, AC, NP contributed data analysis; KH, AJE, MM, SdeL, RGP contributed data collection and provision; AJE, MZ, AC, RGP contributed data interpretation; PJB, XSZ, EvL, AMP, MB, DDA contributed model development and intellectual property. The author(s) read and approved the final manuscript.

### Funding

This work was supported by the Medical Research Council (Unit programme number MC UU 00002/11), the National Institute for Health Research (NIHR) Health Protection Research Units (HPRU) in Respiratory Infections, both in partnership with Public Health England. Neither of the funding bodies had any role in setting up the case study, in the curation, interpretation or analysis of data, or in the writing of the manuscript.

### Availability of data and materials

All datasets used are publicly available. Requests for access should be directed to: Richard.Pebody@phe.gov.uk. Computer code used to carry out the analysis of the SPC model is developed at <https://gitlab.com/pjbirrell/real-time-mcmc>. Code for the implementation of the strain-specific model can be found at [https://github.com/anastasiachtz/COMMAND\\_stan/blob/master/MultiStrainStan.Rmd](https://github.com/anastasiachtz/COMMAND_stan/blob/master/MultiStrainStan.Rmd).

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge Institute of Public Health, Robinson Way, Cambridge, CB2 0SR, UK. <sup>2</sup>National Infection Service, Public Health England, 61 Colindale Avenue, NW9 5EQ, London, UK. <sup>3</sup>Virus Reference Department, National Infection Service, Public Health England, 61 Colindale Avenue, NW9 5EQ, London, UK. <sup>4</sup>Real-time Syndromic Surveillance Team, Field Service, National Infection Service, Public Health England, 5 St Philip's Place, B3 2PW, Birmingham, UK. <sup>5</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, OX2 6GG, Oxford, UK. <sup>6</sup>Royal College of General Practitioners Research and Surveillance Centre, 30 Euston Square, NW1 2FB, London, UK. <sup>7</sup>Centre for Mathematical Modelling of Infectious Disease and Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, WC1H 9SH, UK. <sup>8</sup>MRC Centre for Global Infectious Disease Analysis, Imperial College Faculty of Medicine, London SW7 2AZ, UK.

Received: 3 October 2019 Accepted: 4 March 2020

Published online: 15 April 2020

### References

- Cromer D, Van Hoek AJ, Jit M, Edmunds WJ, Fleming D, Miller E. The burden of influenza in England by age and clinical risk group: A statistical analysis to inform vaccine policy. *J Infect*. 2014;68(4):363–71. Available from: <http://dx.doi.org/10.1016/j.jinf.2013.11.013>.
- Lafond KE, Nair H, Rasooly MH, Valente F, Booy R, Rahman M, et al. Global Role and Burden of Influenza in Pediatric Respiratory Hospitalizations, 1982–2012: A Systematic Analysis. *PLoS Med*. 2016;13(3):e1001977. Available from: <https://dx.plos.org/10.1371/journal.pmed.1001977>.
- Iuliano AD, Roguski KM, Chang HH, Muscatello DJ, Palekar R, Tempia S, et al. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *The Lancet*. 2018;391(10127):1285–300. Available from: <https://www.sciencedirect.com/science/article/pii/S0140673617332932>.
- Dawson P, Gailis R, Meehan A. Detecting disease outbreaks using a combined Bayesian network and particle filter approach. *J Theor Biol*. 2015;370:171–83. Available from: <http://dx.doi.org/10.1016/j.jtbi.2015.01.023>.
- Biggerstaff M, Johansson M, Alper D, Brooks LC, Chakraborty P, Farrow DC, et al. Results from the second year of a collaborative effort to forecast influenza seasons in the United States. *Epidemics*. 2018. Available from: <http://dx.doi.org/10.1016/j.epidem.2018.02.003>.
- Kucharski AJ, Baguelin M. The role of human immunity and social behavior in shaping influenza evolution. *PLoS Pathog*. 2017;13(8):e1006432. Available from: <https://dx.plos.org/10.1371/journal.ppat.1006432>.
- Biggerstaff M, Alper D, Dredze M, Fox S, Chun-Hai I, Hickmann KS, et al. Results from the centers for disease control and prevention's predict the 2013–2014 Influenza Season Challenge. *BMC Infect Dis*. 2016;16:. Available from: <http://view.ncbi.nlm.nih.gov/pubmed/27449080>.
- Cabinet Office. National Risk Register of Civil Emergencies. 2015. Available from: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/419549/20150331\\_2015-NRR-WA\\_Final.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/419549/20150331_2015-NRR-WA_Final.pdf). Accessed 1 Apr 2020.
- Hall IM, Gani R, Hughes HE, Leach S. Real-time epidemic forecasting for pandemic influenza. *Epidemiol Infect*. 2007;135:372–85.
- Baguelin M, Van Hoek AJ, Flasche S, White PJ, Edmunds WJ. Vaccination against pandemic influenza A/H1N1v in England: A real-time economic evaluation. *Vaccine*. 2010;28(12):2370–84.
- Birrell PJ, Pebody RG, Charlett A, Zhang XS, De Angelis D. Real-time modelling of a pandemic influenza outbreak. *Health Technol Assess*. 2017;21(58):. <https://doi.org/10.3310/hta21580>.
- Department of Health. Australian Influenza Surveillance Report: 2017, Season Summary. Aust Gov. 2017. [http://www.health.gov.au/internet/main/publishing.nsf/Content/097F15A91C05FBE7CA2581E20017F09E/\\$File/2017-season-summary-22112017.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/097F15A91C05FBE7CA2581E20017F09E/$File/2017-season-summary-22112017.pdf). Accessed 27 Aug 2019.
- Public Health, England. Surveillance of influenza and other respiratory viruses in the UK: Winter 2017 to 2018. *Publ Health Engl*. 2018. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/740606/Surveillance\\_of\\_influenza\\_and\\_other\\_respiratory\\_viruses\\_in\\_the\\_UK\\_2017\\_to\\_2018.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/740606/Surveillance_of_influenza_and_other_respiratory_viruses_in_the_UK_2017_to_2018.pdf). Accessed 27 Aug 2019.
- Fleming DM, Elliot AJ. Lessons from 40 years' surveillance of influenza in England and Wales. *Epidemiol Infect*. 2008;136(7):866–75. Available from: [https://www.cambridge.org/core/product/identifier/S0950268807009910/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0950268807009910/type/journal_article).
- Boddington NL, Verlander NQ, Pebody RG. Developing a system to estimate the severity of influenza infection in England: Findings from a hospital-based surveillance system between 2010/2011 and 2014/2015. *Epidemiol Infect*. 2017;145(7):1461–70. Available from: [https://www.cambridge.org/core/product/identifier/S095026881700005X/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S095026881700005X/type/journal_article).
- Morbey RA, Elliot AJ, Harcourt S, Smith S, de Lusignan S, Pebody R, et al. Estimating the burden on general practitioner services in England from increases in respiratory disease associated with seasonal respiratory pathogen activity. *Epidemiol Infect*. 2018;146:1389–96.
- Hoschler K, Thompson C, Andrews N, Galiano M, Pebody R, Ellis J, et al. Seroprevalence of influenza A(H1N1) pdm09 virus antibody, England, 2010 and 2011. *Emerg Infect Dis*. 2012;18(11):1894–7. Available from: [http://wwwnc.cdc.gov/eid/article/18/11/12-0720\\_article.htm](http://wwwnc.cdc.gov/eid/article/18/11/12-0720_article.htm).
- de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, et al. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. *Br J Gen Pract J R Coll Gen Pract*. 2017;67(663):440–1. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28963401>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5604796>.
- Public Health England. Sources of UK flu data: influenza surveillance in the UK. 2014. Available from: <https://www.gov.uk/guidance/sources-of-uk-flu-data-influenza-surveillance-in-the-uk>. Accessed 1 Apr 2020.
- Health Protection Agency. UK Severe Influenza Surveillance System (USISS): Protocol for sentinel Acute NHS Trusts 2011–12. 2014. Available

from: [http://web.archive.nationalarchives.gov.uk/20140714043432/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb\\_C/1317132396215](http://web.archive.nationalarchives.gov.uk/20140714043432/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317132396215). Accessed 1 Apr 2020.

21. Birrell PJ, Ketsetzis G, Gay NG, Cooper BS, Presanis AM, Harris RJ, et al. Bayesian modelling to unmask and predict the influenza A/H1N1pdm dynamics in London. *Proc Natl Acad Sci*. 2011;108(45):18238–43.
22. Birrell PJ, Zhang XSS, Pebody RG, Gay NJ, De Angelis D. Reconstructing a spatially heterogeneous epidemic: Characterising the geographic spread of 2009 A/H1N1pdm infection in England. *Sci Rep*. 2016;6. Available from: <http://dx.doi.org/10.1038/srep29004>.
23. Held L, Hofmann M, Höhle M, Schmid V. A two-component model for counts of infectious diseases. *Biostatistics*. 2006;7(3):422–37. Available from: <http://dx.doi.org/10.1093/biostatistics/kxj016>.
24. Chatzilena A, van Leeuwen E, Ratmann O, Baguelin M, Demiris N. Contemporary statistical inference for infectious disease models using Stan. *arXiv preprint*. 2019. Available from: <http://arxiv.org/abs/1903.00423>.
25. Corbella A, Zhang XS, Birrell PJ, Boddington N, Pebody RG, Presanis AM, et al. Exploiting routinely collected severe case data to monitor and predict influenza outbreaks. *BMC Publ Health*. 2018;18(1):. Available from: <http://dx.doi.org/10.1186/s12889-018-5671-7>.
26. De Luca G, Van Kerckhove K, Coletti P, Poletto C, Bossuyt N, Hens N, et al. The impact of regular school closure on seasonal influenza epidemics: a data-driven spatial transmission model for Belgium. *BMC Infect Dis*. 2018;18(1):. Available from: <http://dx.doi.org/10.1186/s12879-017-2934-3>.
27. Czado C, Gneiting T, Held L. Predictive Model Assessment for Count Data. *Biometrics*. 2009;65(4):1254–61. Available from: <http://dx.doi.org/10.1111/j.1541-0420.2009.01191.x>.
28. Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing Optimal Target Populations for Influenza Vaccination Programmes: An Evidence Synthesis and Modelling Study. *PLoS Med*. 2013;10(10):e1001527. Available from: <https://dx.plos.org/10.1371/journal.pmed.1001527>.
29. Pebody RG, Zambon M, Ramsay M. Flu vaccines: an annual challenge. *BMJ (Clin Res Ed)*. 2018;361:k2705. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29950432>.
30. Metcalf CJE, Farrar J, Cutts FT, Basta NE, Graham AL, Lessler J, et al. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *The Lancet*. 2016;388(10045):728–30. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0140673616301647>.
31. de Lusignan S, Correa A. Opportunities and challenges of a World Serum Bank. *Lancet*. 2017;389(10066):250–1. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28118910>.
32. de Lusignan S, Borrow R, Tripathy M, Linley E, Zambon M, Hoschler K, et al. Serological surveillance of influenza in an English sentinel network: pilot study protocol. *BMJ Open*. 2019;9(3):e024285. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30852535http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6429844>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

