BMC Public Health

CrossMark

Exploiting routinely collected severe case data to monitor and predict influenza outbreaks

Alice Corbella[1*†] ⓘ, Xu-Sheng Zhang[2†], Paul J. Birrell[1], Nicki Boddington[2], Richard G. Pebody[2], Anne M. Presanis[2] and Daniela De Angelis[1,2]

Abstract

Background: Influenza remains a significant burden on health systems. Effective responses rely on the timely understanding of the magnitude and the evolution of an outbreak. For monitoring purposes, data on severe cases of influenza in England are reported weekly to Public Health England. These data are both readily available and have the potential to provide valuable information to estimate and predict the key transmission features of seasonal and pandemic influenza.

Methods: We propose an epidemic model that links the underlying unobserved influenza transmission process to data on severe influenza cases. Within a Bayesian framework, we infer retrospectively the parameters of the epidemic model for each seasonal outbreak from 2012 to 2015, including: the effective reproduction number; the initial susceptibility; the probability of admission to intensive care given infection; and the effect of school closure on transmission. The model is also implemented in real time to assess whether early forecasting of the number of admissions to intensive care is possible.

Results: Our model of admissions data allows reconstruction of the underlying transmission dynamics revealing: increased transmission during the season 2013/14 and a noticeable effect of the Christmas school holiday on disease spread during seasons 2012/13 and 2014/15. When information on the initial immunity of the population is available, forecasts of the number of admissions to intensive care can be substantially improved.

Conclusion: Readily available severe case data can be effectively used to estimate epidemiological characteristics and to predict the evolution of an epidemic, crucially allowing real-time monitoring of the transmission and severity of the outbreak.

Keywords: Epidemic monitoring, Bayesian inference, Epidemic models, Influenza, Reproduction number, Severe cases

Background

Recent annual epidemics of influenza have resulted in about 3 to 5 million cases of severe illness each season worldwide [1]. Historically, influenza has always placed a large burden on many national health systems [2], particularly as a result of severe cases in the most at risk groups [3] (e.g. elderly [4], children and people with underlying chronic medical conditions [5], persons living in deprived areas [6]).

Measures of different characteristics of an outbreak, whether from seasonal or a newly emergent strain, are crucial to understand the healthcare burden and plan appropriate response measures. For seasonal influenza, retrospective knowledge of severity and transmissibility provides a valuable baseline measure against which to compare the severity and transmissibility of future pandemics. Prospectively, predictions of the likely extent of transmission and the resulting number of severe cases

*Correspondence: alice.corbella@mrc-bsu.cam.ac.uk
†Alice Corbella and Xu-Sheng Zhang contributed equally to this work.
[1]Medical Research Council, Biostatistics Unit - University of Cambridge, School of Clinical Medicine, Cambridge, UK
Full list of author information is available at the end of the article

Corbella *et al. BMC Public Health* (2018) 18:790

Page 2 of 11

are crucial to anticipate demands on health care facilities (e.g. number of beds in hospital) for each season. These timely predictions are even more crucial to inform prompt targeted responses in the event of a new emerging strain with the potential to cause a pandemic [7].

Epidemic models are increasingly used to understand the effect of particular interventions including: vaccination policies [8]; school closures to reduce transmission in a pandemic [9–11]; reinforced use of antiviral drugs [12]; or changes in hospital management policies.

These models are generally applied to data, such as General Practitioner (GP) consultations for influenza-like illness (ILI) [8, 13] or health-related online queries [14], which are only loosely related to the actual burden and are characterized by highly volatile noise.

By contrast, more specific timely data on a sample of confirmed cases (e.g. confirmed influenza hospitalizations) might be collected routinely by national health systems. An example of these data is the UK Severe Influenza Surveillance System (USISS) [15] that records counts of the weekly Intensive Care Unit (ICU) and High Dependence Unit (HDU) admissions and deaths with confirmed influenza in all hospital trusts in England.

Recently, and in the context of a pandemic, some attention has been paid to estimating and predicting pandemic transmission from routinely collected confirmed-case data [16]. This has entailed the development of a very complicated model which is difficult to use in a seasonal monitoring setting (when less effort is placed on data collection) with a prediction goal. Here we explore a much simpler model to be applied to seasonal influenza, and possibly during a pandemic, relying only on simpler data on severe cases alone, which are timely available. We therefore investigate if data collected through USISS can characterise both seasonal and pandemic epidemics, aiming to achieve both the estimation and the prediction goal.

We formulate an epidemic model that links the available USISS data to the underlying unobserved dynamics of influenza in the UK. The model parameters are inferred using data from the seasonal epidemics in 2012-2015, to obtain nation-level estimates of transmission, as measured by R_n, the average number of new cases generated by an infectious individual in a partially immune population, and severity, as measured by the probability of ICU admission given infection.

Additionally, to assess the predictive power of the model, we perform analyses at different dates within each season. Finally, we study what would happen in the event of a pandemic, when the USISS surveillance scheme would be upgraded to collect more information.

Methods

Data

Following the 2009 pandemic, the World Health Organization (WHO) declared the beginning of a post-pandemic phase [17], encouraging national public health agencies to establish hospital-based surveillance systems to monitor the epidemiology of severe influenza. In response to these guidelines, and to understand the baseline epidemiology of severe influenza, the UK developed a surveillance system to monitor severe cases of influenza, the USISS [18, 19]. After a pilot phase in 2010/11, USISS has run for each influenza season, providing data on laboratory-confirmed ICU/HDU influenza cases and on laboratory-confirmed hospitalized cases.
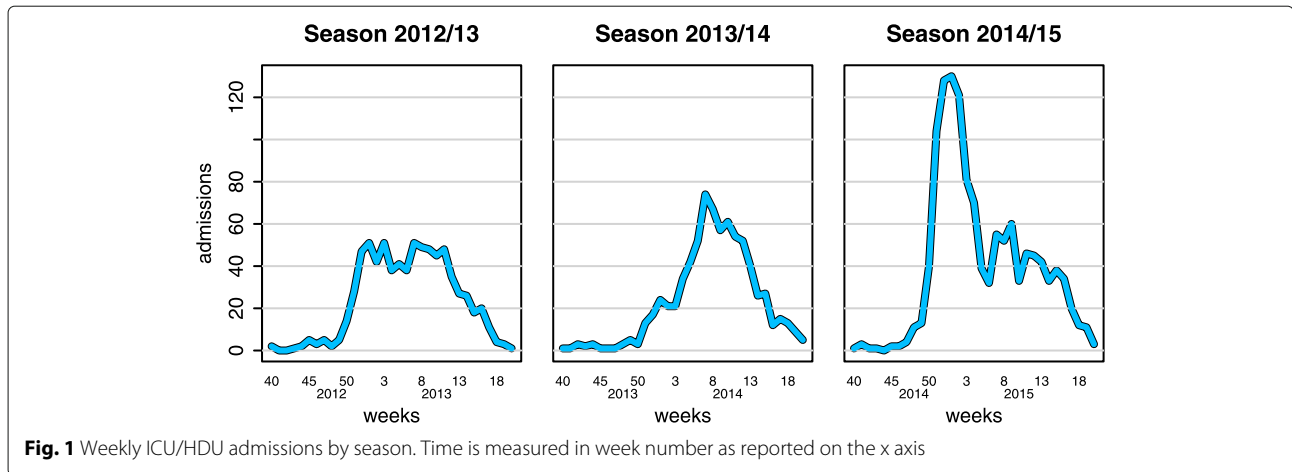
According to the USISS protocol [18], all National Health Service (NHS) trusts report the weekly number of laboratory-confirmed influenza cases admitted to ICU/HDU and the number of confirmed influenza deaths in ICU/HDU via a web tool. An ICU/HDU case is defined as a person who is admitted to ICU/HDU and has a laboratory-confirmed influenza A (including H1, H3 or novel) or B infection.

USISS runs annually from week 40 to week 20 of the following year but, in the event of a pandemic, it can be activated out of this window and will collect the same data at all levels of care, not only ICU/HDU.

Data are available by age group and influenza type/subtype. However, when stratified by both, as well as week, many zero counts are observed. We therefore consider the total ICU/HDU admissions by week only (Fig. 1). Each season between 2012 and 2015 is shown, with each epidemic varying substantially across seasons. In the 2012/13 season, mainly characterized by Influenza B and Influenza A(H3N2) outbreaks, the number of admissions peaks early, maintaining this plateau for several months [20]. In 2013/14, when the predominant strain was A(H1N1), the time series displays a smoother increase, a well localized peak and a subsequent regular decrease [21]. Lastly, in 2014/15, the number of ICU admissions peaks earlier and has a dramatic drop at the beginning of the new year, which is followed by a smaller wave resulting in a time series characterized by a double peak. During this season, Influenza A(H3N2) was the predominant virus circulating and the total number of ICU admissions was higher; this strain is well-known to lead to more severe outcomes, particularly in the elderly [22].

Additional sources of information

In addition to the mandatory scheme, a subgroup of NHS trusts in England is recruited every year to participate in the USISS sentinel scheme [19, 23], which reports weekly numbers of laboratory-confirmed influenza cases hospitalised at all levels of
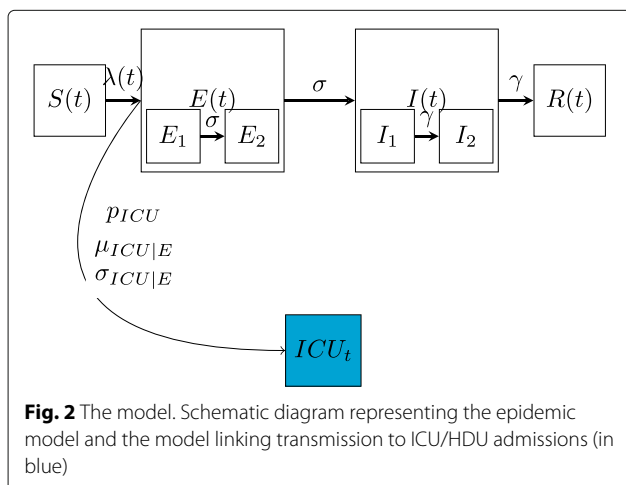
Corbella *et al. BMC Public Health* (2018) 18:790

Page 3 of 11

Fig. 1 Weekly ICU/HDU admissions by season. Time is measured in week number as reported on the x axis

care. From this scheme, individual-level data on all ICU/HDU admissions (until season 2012/13) or on hospital admissions in the young ($\leq$ 17 years old) population (from season 2013/14 onwards) are available, including clinical details such as date of symptom onset, of hospital and ICU admission, and date of discharge from ICU.

These data provide useful information on the process between influenza infection and ICU admission (e.g. the time elapsing from symptom onset to ICU admission). Further information on this process (e.g. proportion of symptomatic cases) can be found in the existing literature about the incubation period of influenza [24] and the hospitalization fatality rate [25].

Model

We used an epidemic model (Fig. 2) to describe the spread of influenza in England [26]. We assumed that the population changes according to a deterministic model in continuous time. Time is measured in days and denoted by $t \geq 0$.

Fig. 2 The model. Schematic diagram representing the epidemic model and the model linking transmission to ICU/HDU admissions (in blue)

The population is divided according to health status into four compartments: susceptible (S), exposed (E), infectious (I) and removed (R). The E and I compartment are further divided into two (E_1, E_2 and I_1, I_2, respectively) so that the waiting times in the E and I states are distributed according to gamma rather than exponential distributions [27]. In the formulas below, the letters S, E_1, E_2, I_1, I_2, R denote the number of people in each compartment. The total size of the population is fixed over every season and denoted by N. The change of compartment is determined by the transition rates: $\lambda(t), \sigma$ and γ explained below.

The infection rate $\lambda(t)$ is proportional to the proportion of people in the infectious compartment at t, $\frac{I_1(t)+I_2(t)}{N}$ and a time varying transmission rate $\beta(t)$:

$$\lambda(t) = \beta(t)\frac{I_1(t) + I_2(t)}{N}. \tag{1}$$

$\beta(t)$ is a function of time and it allows for a scaling factor $\kappa \in (0, 2]$ that expresses the change due to school closure applied to the transmission rate during school opening β_0 [10] as reported in Eq. 2.

$$\beta(t) = \begin{cases} \kappa \cdot \beta_0, & t \in \text{ school holidays} \\ \beta_0, & \text{otherwise.} \end{cases} \tag{2}$$

The transition rates σ and γ are related to the mean latent period, d_L, and the mean infectious period, d_I, by:

$$\sigma = 2/d_L, \qquad \gamma = 2/d_I \tag{3}$$

The system of differential equations that defines the epidemic model is reported in Eq. 4.

Corbella *et al. BMC Public Health* (2018) 18:790

Page 4 of 11

$$\frac{dS}{dt} = -\lambda(t) \cdot S$$

$$\frac{dE_1}{dt} = \lambda(t) \cdot S - \sigma \cdot E_1$$

$$\frac{dE_2}{dt} = \sigma \cdot E_1 - \sigma \cdot E_2$$

$$\frac{dI_1}{dt} = \sigma \cdot E_2 - \lambda \cdot I_1 \qquad (4)$$

$$\frac{dI_2}{dt} = \lambda \cdot I_1 - \lambda \cdot I_2$$

$$\frac{dR}{dt} = \lambda \cdot I_2$$

Here we have assumed homogeneous mixing among contacts (i.e. people are all equally likely to meet, irrespective of their age class and residence, for example).

This transmission model is linked to the data on ICU admissions through an observational model that defines the time elapsing from infection to ICU admission and the probability of ICU admission conditional on infection.

Denote with $f_{ICU|I}(w)$ the probability that w weeks elapse from infection to ICU admission, and with p_{ICU} the probability of ICU admission given infection. We can link μ_w, the average number of ICU admissions during week w, to the weekly new infections in the previous weeks via a convolution:

$$\mu_w = \sum_{v=0}^{w} f_{ICU|I}(w - v) \cdot \Delta I_v p_{ICU} \qquad (5)$$

where $\Delta I_w = (S(w-7) - S(w)) \cdot N$ is the count of the new infections during week w.

To formulate the likelihood of the data, we assumed that the observed number of ICU admissions is the realisation of a Negative Binomial random variable centred on μ_w with over dispersion parameter η:

$$ICU_w \sim \text{NegBin}(\mu_w, \eta), \qquad (6)$$

i.e ICU_w has density function:

$$f(ICU_w = x) = \frac{\Gamma(x + r_w)}{\Gamma(x)\Gamma(x + r_w)} \left(\frac{1}{\eta}\right)^{r_w} \left(1 - \frac{1}{\eta}\right)^{x} \quad (7)$$

with $r_w = \frac{\mu_w}{\eta - 1}$.

The Additional file 1 contains the full specification of the transmission model, its re-parametrization and full derivation of $f_{ICU|I}(w)$.

Parameter estimation

To define the epidemic we need to estimate or set both the transitions rate parameters (i.e. $\beta, \kappa, \sigma, \gamma$) and the initial state of the epidemic (i.e. $S(0), E_1(0), E_2(0), I_1(0), I_2(0), R(0)$).

The epidemic model can be re-parametrized [27] and a number of quantities may be defined, including: π, the initial proportion of non-immune people; $I^{tot}(0) = (I_1(0) +$

$I_2(0))$, the total number of infectious people at $t = 0$; the basic reproduction number R_0 that is the average number of successful transmissions per infectious person in a fully susceptible population; and the effective reproduction number R_n that is the average number of successful transmissions per infectious person in a partially susceptible population. All these parameters are useful under a health-policy perspective.

The parameters σ and γ are assumed known from previous studies [13, 24], as they can be inferred only with detailed information at the individual level. Likewise, the population size N is assumed known and fixed to the values estimated by the Office of National Statistics (ONS) [28].

We used a Bayesian approach to draw inference on the other parameters. Bayesian inference consists in summarizing prior information on a general parameter θ in a distribution $\pi(\theta)$ and updating it with the information deriving from a set of data x, contained in its likelihood $\mathcal{L}(\theta|x)$, to derive the posterior distribution:

$$p(\theta|x) \propto \pi(\theta) \cdot \mathcal{L}(\theta|x). \qquad (8)$$

We considered two scenarios. In the first one we assumed we have no prior information on the values of the parameters except for lower and upper bounds, hence the prior distributions on all the parameters are non-informative (see Additional file 1). Table 1 lists the lower and upper limits of some transformations of the parameters and the values assumed known in this scenario.

Table 1 Prior distributions of the parameters in the non-informative scenario

Unknown parameters		Lower limit	Upper limit
Susceptibility	π	0	1
Initial number of infectious	$I^{tot}(0) = (I_1(0) + I_2(0))$	0	10000
Transmission rate	β	0	1.12
Over-dispersion	η	1	100
P of ICU admission given infection	p_{ICU}	0	1
Scaling factor for school closure	κ	0	2
Parameters assumed known		Value	
Rate of becoming infectious	σ	1	
Rate of recovery	γ	0.5797	
Population of 2012/13	$N_{2012/13}$	53,679,750	
Population of 2013/14	$N_{2013/14}$	54,091,200	
Population of 2014/15	$N_{2014/15}$	54,551,450	

Corbella *et al. BMC Public Health* (2018) 18:790

Page 5 of 11

In the second scenario we used sero-prevalence data from the 2010/11 season [29] to formulate a prior distribution for the initial susceptibility π. The use of sero-prevalence data to describe the immunity of a population could be debatable, since the results may be extendible only to seasons with similar predominant strains circulating. Here, sero-samples were taken during an H1 predominant season: this sub-type was prevalent also in the 2012/13 season, but not in 2014/15. However, combining this prior with the data allows us to test how much prior knowledge is needed to overcome the lack of information about susceptibility from the data. We also derived an informative prior distribution on p_{ICU} by combining estimates of the probability of hospitalization given infection from a previous severity study [25] with estimates of the probability of ICU/HDU admission given hospitalization from the aggregate data of the USISS sentinel scheme. Table 2 lists the prior distributions of the two parameters that change in the informative scenario. The remaining parameters are again assumed to be uniformly distributed.

Analyses

For both the prior settings we performed two types of analysis: firstly we considered all the data reported in Fig. 1 and we analysed them retrospectively. Secondly, to assess the predictive ability of our model, we performed estimation and forecasting assuming only an initial portion of the data are available. We used the data up to week w as a training dataset to estimate the parameters. Then we predicted the evolution of the epidemic after week w, based on the estimates from the training dataset. We tested the following prediction time points: $w = 3, 8, 13,$ and 18 from the beginning of the new year.

To approximate the posterior distribution, we used a Metropolis Hastings block updated sampling algorithm [30], coded using the R programming language [31]. The system of differential Eq. (4) was solved using the R package deSolve [32]. Details on the algorithm are available in the Additional file 1 and the code is available at http://www.mrc-bsu.cam.ac.uk/software/miscellaneous-software/.

Results

Retrospective analysis

The retrospective analysis of the data was first performed in the uninformative scenario. The resulting posterior
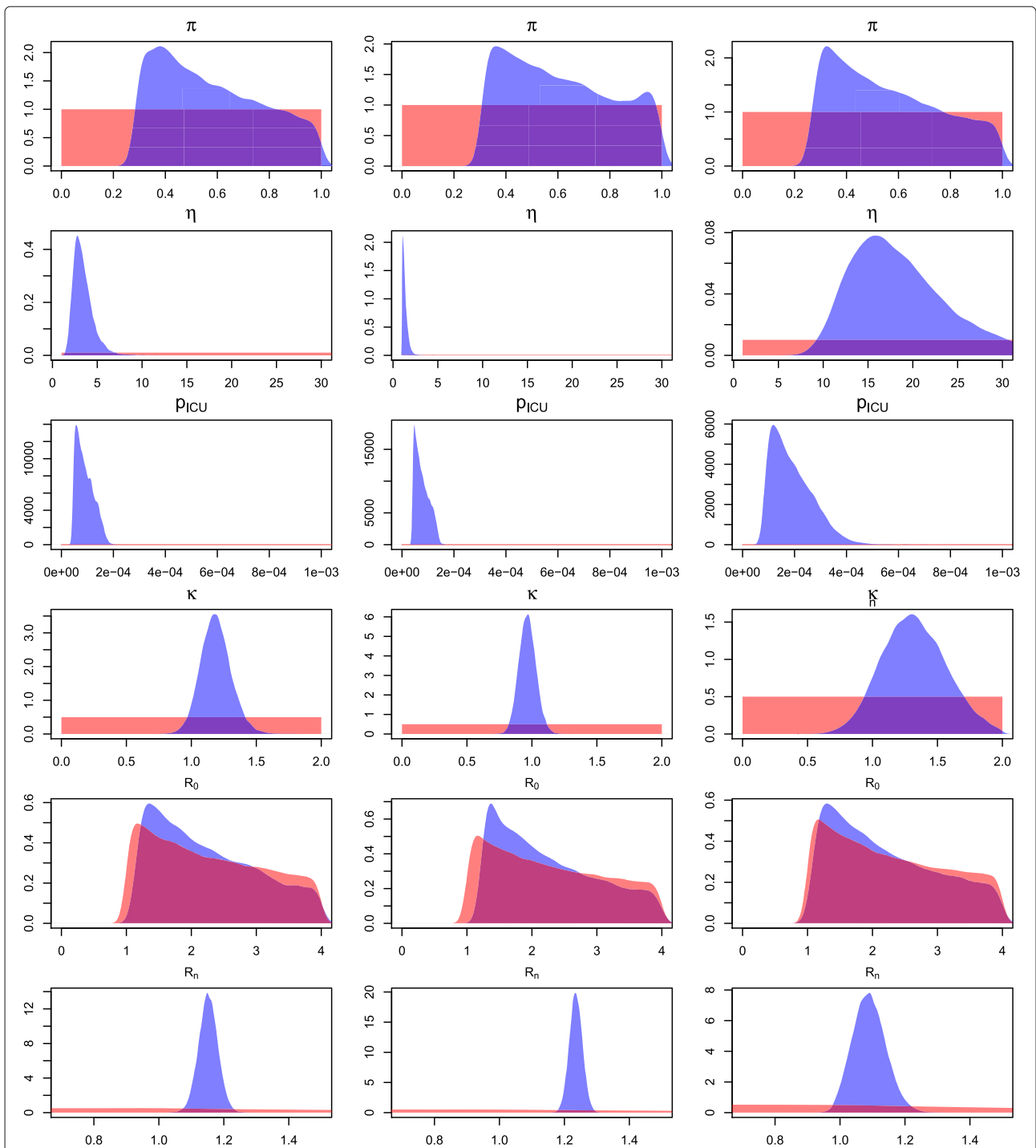
distributions are displayed in Fig. 3 with the posterior median and 95% Credible Intervals (CrI)s of some of the parameters reported in Table 3. Note that the posterior distribution of the basic reproduction number R_0 is almost identical to the prior. This is due to the fact that the information contained in the data is not sufficient to determine separately the values of the parameters describing both the initial immunity and the transmission rate. For the same reason the posterior distribution of the parameter π doesn't change significantly from its prior, only excluding those small values that would completely prevent an epidemic to take place. This problem is explored in detail in the Additional file 1.

Data are much more informative about parameters η, p_{ICU} and κ. The highly variable behaviour of the ICU admissions count in season 2014/15 is reflected by the over-dispersion parameter η, whose distribution is significantly higher compared to the ones estimated from the 2012/13 and 2013/14 seasons. The range of the probability of going to ICU given infection, p_{ICU}, is always between 0.004 and 0.04%. Its median is higher in season 2014/15, in agreement with the higher severity that was detected during this influenza season [23]. The multiplicative factor κ introduced to allow for a school-closure effect is centred on 1 for season 2013/14 and centred around higher values in the remaining seasons. A possible explanation for this counter-intuitive phenomenon relies on the age distribution of the sample population. Our data have a different distribution compared to the English population [23, 28], with patients over 65 being over represented and children in school years being under represented. The elderly individual perhaps are more likely to meet other potential influenza spreaders (e.g. children) during school closures, particularly over Christmas holiday. It makes sense, therefore, to observe an inverse relationship between school closure and the transmission rate, in contrast to results that might be expected from a more representative sample of the population [10]. However, this piecewise increment in transmission rate may incorporate other time-varying phenomena that affect the force of infection. The Christmas holiday often coincides with the beginning of a colder and more humid period and changes in vapour pressure, that might imply an increasing spread of influenza [33]. Lastly the posterior median of the effective reproduction number R_n is equal to 1.152, 1.235, 1.089 in seasons 2012/13, 2013/14 and 2014/15 respectively.

Although the CrIs of the parameter κ included 1, the posterior probability of it being larger than 1 ($\Pr(\kappa > 1)$) is substantial for two seasons. The introduction of this parameter allows the flexibility needed to represent the specific features of each season. This can be observed in the posterior predictive distribution of the weekly ICU admissions reported in Fig. 4. Specifically in season

Table 2 Prior distributions of the parameters that change in the informative scenario

Parameters		Distribution
Susceptibility	π	$\sim$ LogNorm(log μ = log(0.401), log σ = 0.2) [29]
P of ICU admission given infection	p_{ICU}	$\sim$ LogNorm(log μ = log(0.000239), log σ = 1) [25]

Corbella *et al. BMC Public Health* (2018) 18:790

Page 6 of 11

Fig. 3 Retrospective analysis, uninformative scenario. Prior (red) and posterior (blue) distributions of: the initial susceptibility (π); the over-dispersion parameter (η); the probability of ICU admission given infection (p_{ICU}); the scaling parameter (κ); and the basic and effective reproduction number (R_0 and R_n). The results are derived from season 2012/13 (left column), season 2013/14 (centre) and season 2014/15 (right column)

2012/13 we manage to reproduce the plateau that takes place from the end of the Christmas vacations to the February half term. Regarding instead the double peaking season of 2014/15, the 95% Credible bounds are not narrow, but the timing of the peak of the distribution is

predicted substantially better than in the case of constant infection rate (results not shown). The high variability of the data considered, combined with the constraint of a deterministic model, cause an overall poor fitting of the model to the data of this season. This model does not

Corbella *et al. BMC Public Health* (2018) 18:790

Page 7 of 11

Table 3 Posterior medians and 95% CrIs from the retrospective analysis of the ICU admissions with uninformative priors

Parameter		Season 2012/13 Posterior Me (CrI)	Season 2013/14 Posterior Me (CrI)	Season 2014/15 Posterior Me (CrI)
Susceptibility	π	0.546 (0.297 - 0.969)	0.589 (0.32 - 0.977)	0.531 (0.28 - 0.968)
Initial number of infectious	I^{tot}	4106 (1441 - 11510)	1357 (484 - 3312)	9590 (3053 - 28493)
Transmission rate	β	0.611 (0.344 - 1.126)	0.608 (0.367 - 1.118)	0.596 (0.324 - 1.119)
Over-dispersion	η	3.204 (1.888 - 6.101)	1.25 (1.011 - 2.096)	17.925 (10.412 - 35.812)
P of ICU given infection	$p_{ICU}(\cdot 10^4)$	0.841 (0.458 - 1.614)	0.713 (0.419 - 1.338)	1.749 (0.848- 3.745)
Scaling factor for school closure	κ	1.185 (0.971 - 1.434)	0.965 (0.841 - 1.1)	1.313 (0.866 - 1.824)
Effective reproduction number	R_n	1.152 (1.093 - 1.211)	1.235 (1.196 - 1.275)	1.089 (0.997 - 1.195)

allow precise inference both of the parameters and of the predictions.

The same analysis was performed in the second scenario, i.e. allowing informative priors on the susceptibility π and on p_{ICU} as defined in Table 2. The introduction of these prior distributions compensates for the lack of information, allowing the identification of π and improving the precision of the posterior distribution of p_{ICU}. This affects also other parameters such as β and R_0. However, their posterior distributions are driven by the prior distributions alone, and they do not learn from the data. In terms of fit there was no improvement. Results are reported in the Additional file 1.
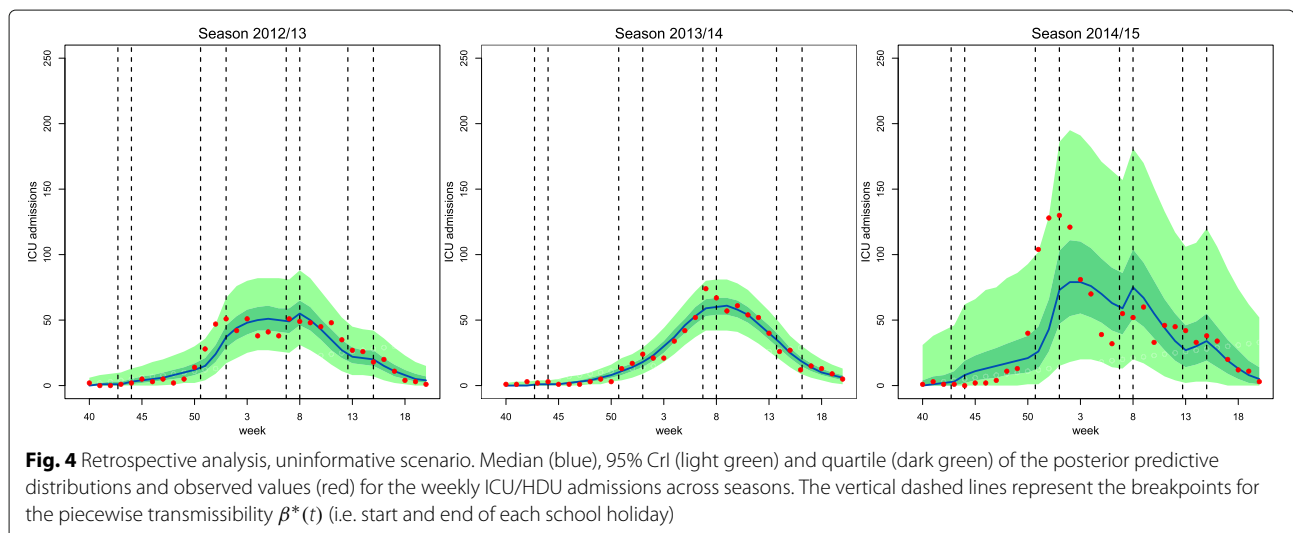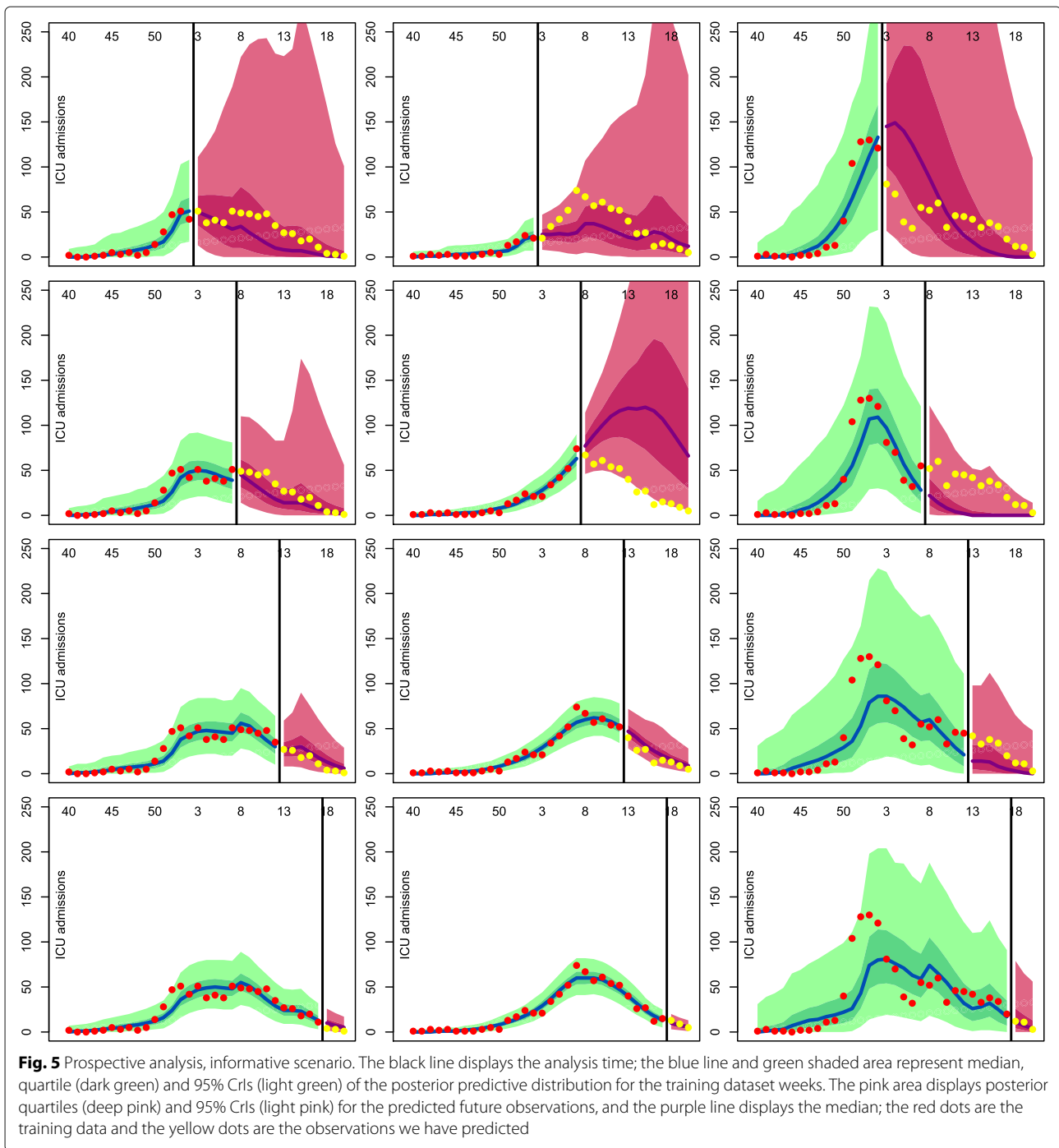
Prediction

The prospective analysis of the data in the uninformative scenario resulted in very wide predictions of the future dynamics, therefore we assumed the informative priors reported in Table 2. The performance of the model at different times is plotted in Fig. 5 for each season.

Season 2013/14, despite displaying the most regular data, is the most difficult to predict: the well-defined initial growth biases the predictions towards a major outbreak. This leads to the median and the credible intervals of the posterior predictive distribution over-estimating the data until mid-march (week 13 from the beginning of the year). For the other two seasons, the median predicted weekly ICU admissions is always very close to the data points, but the credible intervals narrow to reasonable bounds only towards the end of February (week 8 from the beginning of the year).

Prediction is challenging, as demonstrated by the precision of the predictions. For example, the 95% CrI of the predicted number of ICU admissions 3 weeks in advance, when the epidemic is still taking off (i.e. at the third week of January) is as wide as 138 for season 2012/2013 (from 2 to 140 ICU admissions), 52 for season 2013/2014 (from 6 to 58 ICU admissions) and 473 for season 2014/2015 (from 11 to 484 ICU admissions). Due to the different sizes of the epidemics, the coefficient of variation (i.e. the ratio of the posterior standard deviation to the posterior mean) can be used to compare them: it is equal to 0.751 for season 2012/13, 0.491 for season 2013/14, and 0.742 for season 2014/15, highlighting that predictions prediction precision increases when the epidemic is smaller and less over-dispersed.

Fig. 4 Retrospective analysis, uninformative scenario. Median (blue), 95% CrI (light green) and quartile (dark green) of the posterior predictive distributions and observed values (red) for the weekly ICU/HDU admissions across seasons. The vertical dashed lines represent the breakpoints for the piecewise transmissibility $\beta^*(t)$ (i.e. start and end of each school holiday)

Corbella *et al. BMC Public Health* (2018) 18:790

Page 8 of 11

Fig. 5 Prospective analysis, informative scenario. The black line displays the analysis time; the blue line and green shaded area represent median, quartile (dark green) and 95% CrIs (light green) of the posterior predictive distribution for the training dataset weeks. The pink area displays posterior quartiles (deep pink) and 95% CrIs (light pink) for the predicted future observations, and the purple line displays the median; the red dots are the training data and the yellow dots are the observations we have predicted

In spite of the simplicity of our model, the flexibility introduced by the parameter κ allows for the correction "on the fly" of the prediction, adapting to new peaks (e.g. season 2014/15) or periods of constant influenza circulation (e.g. season 2012/13).

Nonetheless, similarly to most epidemic models attempting predictions [13, 34], results are not useful (i.e. precise enough to determine a health policy response) until after the epidemic has peaked.

Further results

We simulated the weekly count of Hospital admissions in the case of a pandemic and we extended our model enabling the inference of the parameters from these data. Despite the increased number of observations, the model performed very similarly to the case of non-pandemic ICU-counts data. We diagnosed identifiability problems in the uniform prior scenario and predictions were good only when more informative prior distributions

Corbella *et al. BMC Public Health* (2018) 18:790

Page 9 of 11

(on the susceptibility and probability of hospitalization) were included. Results from this analysis are reported in Section 5 of the Additional file 1.

Other analyses performed include: prospective analysis for the uninformative scenario and retrospective analysis within the informative scenario. Results of these analyses are reported in Section 4 of the Additional file 1.

Discussion

In this paper we proposed a model to estimate and predict influenza outbreaks from routinely collected data on admissions to ICU/HDU.

We investigated the performance of the proposed model both on simulated and on real data. By fitting the model to simulated numbers of weekly ICU admissions, we discovered that, even with very vague prior information, we could obtain estimates of some of the main parameters, including the initial infection rate, the probability of going to ICU given infection, the effective reproduction number R_n and the scaling factor for school holidays κ. When we injected information on the distribution of the average immunity $(1 - \pi)$ and on p_{ICU}, estimates of the remaining parameters could be obtained. We were also able to forecast the evolution of the outbreak by analysing the first months of the epidemic using data up to the peak of influenza activity.

The model was applied to real data on the weekly number of ICU admissions from seasons 2012/13, 2013/14 and 2014/15, confirming the performance obtained on the simulated data. The estimated values of the effective reproduction number R_n were similar to those estimated during the past decade of seasonal influenza [8]. A scaling parameter allowed the transmission rate to vary between school and holiday/half-term periods, which resulted in a good fit of the model to the data for most of the seasons considered. A more complete investigation of the temporal variation of the transmission rate might improve the flexibility of our model, and therefore the fit to more anomalous epidemics.

Recently, a similar analysis was performed on the Finnish influenza pandemic of 2009 [16] using a more elaborate model, analysing confirmed data on both hospitalizations and GP consultation. Their inclusion of GP data enhances the performance of the inference. Nevertheless, these data are harder to collect in a larger population (England is almost 10 times more populated than Finland) and out of pandemic emergencies. By contrast, the inference performed through our model is driven by few data, though readily available, even in real time, in seasonal settings. A further advance of the model by [16] is that the transmission parameter is time varying according to a Gaussian Process: this allows an accurate description of the past dynamics but makes prediction infeasible, since this temporal variation cannot be forecast. By contrast, our simple piecewise constant model is able to well forecast the future trend and it includes enough flexibility to describe appropriately the present and the past data.

Our work has also some limitations: firstly, our model is non-age-specific. The assumption of homogeneous mixing across regions and age groups is very strong but this was dictated by the very small sample sizes which did not allow sub-grouping. Secondly, the quality of some estimates and predictions strongly relies on prior information on the proportion of non-immune people. As this information is needed to overcome the lack of identifiability in the parameters, we used sero-prevalence data following the 2010/11 epidemic. This is not likely to be correct for all the three seasons analysed, as the predominant strain circulating was different across seasons. Likewise, the model that describes the time elapsing between infection and ICU admission, is assumed to be fixed and mostly known, but this assumption is not likely to be valid. The other element that defines the observational process, i.e. the probability of ICU admission given infection, is also sensitive to the choice of prior distribution.

Conclusion

The work presented here is a proof of concept of the potential for estimation and prediction of influenza transmission from USISS data. At the same time, the results highlight the need of collecting external data to formulate an appropriate prior distribution on the initial immunity of the population, particularly in the event of a pandemic.

The availability of this information, together with the tool we have provided here, allows to retrospectively infer the epidemic parameters from routinely collected data on severe cases during seasonal outbreaks and to predict the temporal dynamics of new epidemics.

Additional file

Additional file 1: Appendix to the main text, including further explanation of the models and algorithms used, description of the methodological challenges and supplementary results. (PDF 8040 kb)

Corbella *et al. BMC Public Health*   (2018) 18:790

Page 10 of 11

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

Author details
[1]Medical Research Council, Biostatistics Unit - University of Cambridge, School of Clinical Medicine, Cambridge, UK. [2]Centre for Infectious Disease Surveillance and Control, Public Health England, London, UK.

References

1. World Health Organization (WHO). Influenza (seasonal) fact sheet. Available from: http://www.who.int/mediacentre/factsheets/fs211/en/. Accessed Oct 2017.
2. Pitman RJ, Melegaro A, Gelb D, Siddiqui MR, Gay NJ, Edmunds WJ. Assessing the burden of influenza and other respiratory infections in England and Wales. J Infect. 2007;54(6):530–8.
3. Hayward AC, Fragaszy EB, Bermingham A, Wang L, Copas A, Edmunds WJ, et al. Comparative community burden and severity of seasonal and pandemic influenza: Results of the Flu Watch cohort study. Lancet Respir Med. 2014;2(6):445–54.
4. Matias G, Taylor R, Haguinet F, Schuck-Paim C, Lustig R, Shinde V. Estimates of mortality attributable to influenza and RSV in the United States during 1997–2009 by influenza type or subtype, age, cause of death, and risk status. Influenza Other Respir Viruses. 2014;8(5):507–15.
5. Neuzil KM, Wright PF, Mitchel EF, Griffin MR. The burden of influenza illness in children with asthma and other chronic medical conditions. J Pediatr. 2000;137(6):856–64.
6. Zhao H, Harris R, Ellis J, Pebody R. Ethnicity, deprivation and mortality due to 2009 pandemic influenza A (H1N1) in England during the 2009/2010 pandemic and the first post-pandemic season. Epidemiol Infect. 2015;143(16):3375–83.
7. Van Kerkhove MD, Asikainen T, Becker NG, Bjorge S, Desenclos Jc, Santos T, et al. Studies Needed to Address Public Health Challenges of the 2009 H1N1 Influenza Pandemic : Insights from Modeling. PLoS Med. 2010;7(6):1–6.
8. 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):1–19.
9. House T, Baguelin M, Van Hoek AJ, White PJ, Sadique Z, Eames K, et al. Modelling the impact of local reactive school closures on critical care provision during an influenza pandemic. Proc Biol Sci. 2011;278(1719): 2753–60.
10. Te Beest DE, Birrell PJ, Wallinga J, De Angelis D, van Boven M. Joint modelling of serological and hospitalization data reveals that high levels of pre-existing immunity and school holidays shaped the influenza A pandemic of 2009 in the Netherlands. J R Soc Interface. 2015;12(103): 20141244.
11. Vynnycky E, Edmunds WJ. Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures. Epidemiol Infect. 2008;136(2):166–79.
12. Ferguson NM, Cummings DAT, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006;442(7101): 448–52.
13. Birrell PJ, Ketsetzis G, Gay NJ, Cooper BS, Presanis AM, Harris RJ, et al. Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. Proc Natl Acad Sci. 2011;108(45):18238–43.
14. Yang W, Lipsitch M, Shaman J. Inference of seasonal and pandemic influenza transmission dynamics. Proc Natl Acad Sci. 2015;112(9):2723–8.
15. 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 14 Aug 2017.
16. Shubin M, Lebedev A, Lyytikäinen O, Auranen K. Revealing the True Incidence of Pandemic A(H1N1)pdm09 Influenza in Finland during the First Two Seasons - An Analysis Based on a Dynamic Transmission Model. PLoS Comput Biol. 2016;12(3):1–19.
17. World Health Organization (WHO). H1N1 in post-pandemic period. Available from: http://www.who.int/mediacentre/news/statements/2010/h1n1_vpc_20100810/en/. Accessed Oct 2017.
18. Health Protection Agency (HPA). UK Severe Influenza Surveillance System (USISS) Protocol for all NHS Acute Trusts 2011-12; 2011. Available from: http://webarchive.nationalarchives.gov.uk/20140714043422/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317132396353. Accessed Oct 2017.
19. Health Protection Agency (HPA). UK Severe Influenza Surveillance System (USISS) Protocol for sentinel Acute NHS Trusts 2011-12; 2011. Available from: http://webarchive.nationalarchives.gov.uk/20140714043432/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317132396215. Accessed Oct 2017.
20. Public Health England. Surveillance of influenza and other respiratory viruses, including novel respiratory viruses, in the United Kingdom: Winter 2012/13; 2013. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/325217/Annual_flu_report_winter_2012_to_2013.pdf. Accessed 29 Mar 2017.
21. Public Health England. Surveillance of influenza and other respiratory viruses in the United Kingdom: Winter 2013/14; 2014. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/325203/Flu_annual_report_June_2014.pdf. Accessed 29 Mar 2017.
22. Public Health England. Surveillance of influenza and other respiratory viruses in the United Kingdom: winter 2014 to 2015; 2015. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/429617/Annualreport_March2015_ver4.pdf. Accessed 29 Mar 2017.
23. Boddington NL, Verlander NQ, Pebody RG, the UK Severe Influenza Surveillance System Steering Group. 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.
24. Tom BDM, van Hoek AJ, Pebody R, McMenamin J, Robertson C, Catchpole M, et al. Estimating time to onset of swine influenza symptoms after initial novel A(H1N1v) viral infection. Epidemiol Infect. 2011;139(9): 1418–24.
25. Presanis AM, Pebody RG, Birrell PJ, Tom BDM, Green RK, Durnall H, et al. Synthesising evidence to estimate pandemic (2009) A/H1N1 influenza severity in 2009-2011. Ann Appl Stat. 2014;8(4):2378–403.
26. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton New Jersey: Princeton University Press; 2008.
27. Wearing HJ, Rohani P, Keeling MJ. Appropriate models for the management of infectious diseases. PLoS Med. 2005;2(7):0621–7.
28. Office of National Statistics. Population Estimates for UK, England and Wales, Scotland and Northern Ireland; 2012-2015. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/

Corbella *et al. BMC Public Health*   (2018) 18:790

Page 11 of 11

populationestimatesforukenglandandwalesscotlandandnorthernireland.
Accessed 23 June 2016.

29. 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.

30. Robert C, Casella G. Introducing Monte Carlo Methods with R. New York:
Springer Science & Business Media; 2009.

31. R Core Team. R: A Language and Environment for Statistical Computing.
Vienna, Austria; 2017. Available from: https://www.R-project.org/.
Accessed Oct 2017.

32. Soetaert K, Petzoldt T, Setzer RW. Solving differential equations in R:
package deSolve. J Stat Softw. 2010;33(9):33.

33. Lipsitch M, Viboud C. Influenza seasonality: lifting the fog. Proc Natl Acad
Sci. 2009;106(10):3645–6.

34. Ong JBS, cheng Chen MI, Cook AR, Lee HC, Lee VJ, Lin RTP, et al.
Real-time epidemic monitoring and forecasting of H1N1-2009 using
influenza-like illness from general practice and family doctor clinics in
singapore. PLoS One. 2010;5(4):1–11.

35. UW, Department of Biostatistics. Summer School in Statistics and
Infectious Disease Modelling; 2016. Available from: http://www.biostat.
washington.edu/suminst/sismid. Accessed Oct 2017.