IDSA Infectious Diseases Society of America   **hivma** hiv medicine association   OXFORD

Seasonal Influenza Forecasting in Real Time Using the Incidence Decay With Exponential Adjustment Model

Tahmina Nasserie,[1,2] Ashleigh R. Tuite,[6] Lindsay Whitmore,[7] Todd Hatchette,[8,9] Steven J. Drews,[10,11] Adriana Peci,[5] Jeffrey C. Kwong,[2,4,5,6] Dara Friedman,[7] Gary Garber,[5] Jonathan Gubbay,[5] and David N. Fisman[2,3]

[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada; [2]Division of Epidemiology, Dalla Lana School of Public Health, and [3]Division of Infectious Diseases and [4]Department of Family and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada; Institute for Clinical and Evaluative Sciences, Sunnybrook Hospital, Toronto, Ontario, Canada; and [5]Public Health Ontario Toronto, Ontario, Canada; [6]Prevention Policy Modeling Laboratory, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [7]Ottawa Public Health, Ottawa, Ontario, Canada; [8]Department of Pathology and Laboratory Medicine; [9]Dalhousie University, Halifax, Nova Scotia, Canada; Capital District Health Authority Nova Scotia, Halifax, Nova Scotia, Canada; [10]ProvLab Alberta Health Services, Edmonton, Alberta Canada; and [11]Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada

Background. Seasonal influenza epidemics occur frequently. Rapid characterization of seasonal dynamics and forecasting of epidemic peaks and final sizes could help support real-time decision-making related to vaccination and other control measures. Real-time forecasting remains challenging.

Methods. We used the previously described "incidence decay with exponential adjustment" (IDEA) model, a 2-parameter phenomenological model, to evaluate the characteristics of the 2015–2016 influenza season in 4 Canadian jurisdictions: the Provinces of Alberta, Nova Scotia and Ontario, and the City of Ottawa. Model fits were updated weekly with receipt of incident virologically confirmed case counts. Best-fit models were used to project seasonal influenza peaks and epidemic final sizes.

Results. The 2015–2016 influenza season was mild and late-peaking. Parameter estimates generated through fitting were consistent in the 2 largest jurisdictions (Ontario and Alberta) and with pooled data including Nova Scotia counts (R_0 approximately 1.4 for all fits). Lower R_0 estimates were generated in Nova Scotia and Ottawa. Final size projections that made use of complete time series were accurate to within 6% of true final sizes, but final size was using pre-peak data. Projections of epidemic peaks stabilized before the true epidemic peak, but these were persistently early (~2 weeks) relative to the true peak.

Conclusions. A simple, 2-parameter influenza model provided reasonably accurate real-time projections of influenza seasonal dynamics in an atypically late, mild influenza season. Challenges are similar to those seen with more complex forecasting methodologies. Future work includes identification of seasonal characteristics associated with variability in model performance.

Keywords. epidemics; epidemiology; forecasting; influenza; mathematical modeling.

Seasonal epidemics of influenza A and B viruses occur frequently. These viruses infect the respiratory tract and typically cause fever, cough, and muscle and joint pain [1]. In some individuals, particularly young children and the elderly, influenza infection may lead to severe illness or death [1–3]. Seasonal influenza also places a strain on economic productivity due to employee absenteeism [4] and exhaustion of healthcare resources due to increased hospitalization [5].

The ability to rapidly characterize and forecast influenza epidemics using mathematical models could help support evidence-driven, real-time decision-making for influenza preparedness and response. Key activities that could be supported by

rapid characterization and forecasting of epidemic dynamics include optimal allocation of vaccines, hospital resource management, and health communications. Mechanistic approaches such as compartmental models and agent-based models are commonly used to forecast influenza activity [6, 7], but these typically require abundant data on population mixing and immunity patterns to produce relevant insights. The incidence decay with exponential adjustment (IDEA) model is a single-equation approach that requires knowledge of only 2 parameters [8–10]. The IDEA model has had some success in rapidly describing epidemic dynamics and forecasting epidemic peak and end dates based on limited data, and it has been applied to recent outbreaks of cholera [11], Middle East respiratory syndrome (MERS) [8], and Ebola [9].

We sought to apply the IDEA model to seasonal influenza surveillance data from participating Canadian provincial laboratories and a city (Ottawa) health department. The goal of this study was to evaluate the use of this model as a tool for real-time surveillance and projection of seasonal influenza epidemic peaks. In particular, our objectives were to (1) determine whether the IDEA model can rapidly generate accurate estimates of the basic reproduction number for a given influenza

season and (2) use these estimates to generate accurate forecasts of seasonal influenza epidemics, with a particular focus on peak and end dates, and final epidemic size.

METHODS

Data Sources

This study made use of anonymized weekly public health surveillance data from Ottawa Public Health and laboratory surveillance data from the Ontario, Alberta, and Nova Scotia virology laboratories. Laboratory datasets provided information about the total number of tests performed as well as the number of tests that were positive for influenza A virus or B virus. The Ottawa Public Health (city) dataset provided information about the onset date of influenza illness, the requisition date, and the date when the case was reported as test positive by the relevant laboratory.

The Alberta (ProvLab Alberta [PLA]) and Ontario (Public Health Ontario Laboratories [PHOL]) laboratory systems perform testing on 25 000–50 000 influenza specimens in a given season, with substantial variability depending on the severity of the influenza season. ProvLab Alberta performs testing on all specimens submitted for respiratory testing in the province, whereas PHOL performs testing predominantly focuses on respiratory specimens submitted from hospitals and emergency rooms, although community specimens are also tested. It has been estimated that PHOL tests approximately 55% of unique respiratory virus specimens in the Province of Ontario (Jeff Kwong, personal communication, written communication, February 2, 2017). The Microbiology Laboratory of the Queen Elizabeth II Health Centre (Central Zone of the Nova Scotia Health Authority) (QE2) performed testing in Nova Scotia. Unlike the PHOL and PLA systems, QE2 testing is limited to hospital testing and testing related to outbreaks in long-term care facilities.

Model

The IDEA is a single-equation approach that describes epidemic processes in terms of exponential growth and simultaneous decay [10]. In contrast with traditional approaches, IDEA aims to represent factors other than population immunity that are associated with slowing epidemic growth [10]. The IDEA is described by a single equation, which takes the functional form $It = = [R_0/(1+d)^t]^t$. The model describes epidemic processes in terms of exponential growth (a function of the basic reproductive number, R_0) and simultaneous decay (d) brought about by behavioral change, public health interventions, increased immunity in the population, or any other dynamic change that slows disease transmission.

R_0 and d were estimated using a Bayesian pseudo-likelihood approach as described by Bolker [12], with 95% confidence intervals generated assuming a negative binomial distribution [13]. Models were updated each generation as new weekly data became available.

Analyses

Models were fit to aggregate weekly incident case counts reported by the Ontario (PHO), Alberta (PLA), and Nova Scotia (QE2) laboratories and Ottawa Public Health between September 2015 and May 2016, as well as to the summed total incident case counts across all laboratories (but excluding data from Ottawa Public Health, to avoid double counting, because some of these cases are also included in the Ontario dataset).

For simplicity, the mean generation time for influenza infection was assumed to be 3 days [14], meaning that weekly case counts represented cases produced in approximately 2 generations, although this was varied in sensitivity analyses. The onset of influenza season was defined as 2 consecutive weeks in which the number of cases was greater than zero and greater than the number of cases reported in the prior week. The end of influenza season was defined as 2 weeks in a row with zero cases. For prospective analyses, the IDEA model was fit to incident case data to predict epidemic peaks, end dates, and final sizes. Models were updated each generation as new weekly data became available. The start dates of influenza season in each region were November 22, 2015 (Alberta), November 25, 2015 (Ottawa), November 29, 2015 (Ontario), and December 28, 2015 (Nova Scotia).

RESULTS

The 2015–2016 Influenza Season

The number of incident influenza cases reported by all participating laboratories over time is shown in Figure 1. Alberta and Ontario laboratories reported most cases (42.7% and 50.8% of total cases). A high proportion of cases were reported between mid-February and mid-March 2016. Overall, the epidemic appeared to peak near the end of February 2016. Influenza activity in this season started much later than most previous influenza seasons, especially in Nova Scotia, where the season
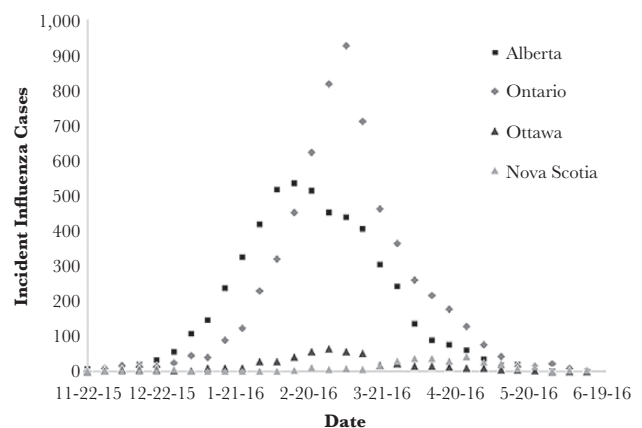

Figure 1. Weekly incidence of influenza cases in Ottawa, Ontario, Alberta, and Nova Scotia. The vertical axis represents the number of influenza cases reported by participating laboratories and Ottawa Public Health, covering the period between week 48 (beginning November 23, 2015) and week 21 (beginning May 23, 2016).

started in late December 2015. Influenza activity continued until late May 2016. As with other Northern Hemisphere geographies, the 2015–2016 influenza season appeared relatively mild [15]. Overall, the model reproduced influenza trends well when fit to the complete time series combined across all jurisdictions (Supplementary Figure 1).

Parameter Estimation

Epidemic dynamics differed slightly in the different regions involved in this study (Supplementary Figures 2–5). For all models, there was a gradual increase in best-fit R_0 and d values as the influenza season wore on. Using all available data for 2015–2016, we noted that the highest R_0 estimates were associated with jurisdictions with the largest case counts (Alberta and Ontario, with R_0 values of approximately 1.4), whereas lower case counts were associated with lower R_0 estimates (Ottawa and Nova Scotia, R_0 values 1.1 to 1.2) (Table 1). Higher R_0 values were associated with higher d values (Table 1). The best-fit R_0 estimates are similar to those often estimated for influenza [16].

When provincial case counts for Alberta, Ontario, and Nova Scotia were summed, R_0 estimates were similar to those seen in Alberta and Ontario (~1.4), with d ~0.005 (again, as seen in estimates from Ontario and Alberta). Sensitivity analyses on influenza generation time resulted in a range of R_0 estimates from 1.2 to 1.5 (Table 1). For models using pooled case counts, parameter estimates were stable by early February (approximately 24 generations) with little change thereafter (Figure 2). There were no sudden surges in either estimated R_0 or d throughout the duration of the outbreak.

Forecasting Final Size and Epidemic Peak

Projections were generated using best-fit parameter estimates. With complete data summed across jurisdictions (ie, at the end of May 2016), model projected case counts were
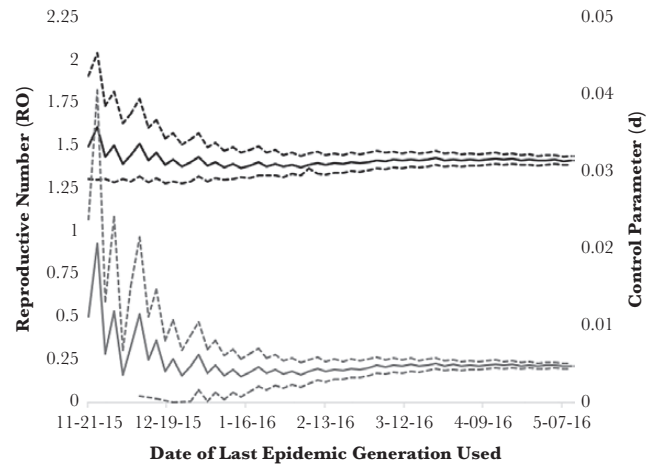

Figure 2. Parameter estimates, summed provincial case counts. Parameter estimates (solid lines; R_0: left y-axis, and d: right y-axis) for summed influenza counts from Ontario, Alberta, and Nova Scotia. Dashed lines represent 95% confidence intervals for parameter estimates. Overall influenza activity had begun in late November 2015; parameter estimates had stabilized by mid-January 2016. The dates on the x-axis represent the date of the most recent influenza data used for model fitting.

similar to observed case counts (11 002 vs 11 686 cases [a 6.2% difference in counts]). Final size projections were fairly stable from mid-February 2016 onwards, but the precision of these estimates improved markedly after early March 2016 (Figure 3). Projections based on earlier parameters tended to overestimate final size (for example, from mid- to late-February 2016, final size estimates ranged between approximately 13 000 to 15 500 cases). Projections were similar using 2- or 4-day generation times. Jurisdiction-level projections of final size (eg, those for Alberta [Supplementary Figure S7] and Ontario [Supplementary Figure S9]) were similar to overall projections but less stable, with persistently high upperbound confidence limits.

Estimates of epidemic peak dates for combined data are shown in Figure 4. Near the beginning of the outbreak, in December 2015, estimates of epidemic peak dates were inaccurate and tended to fluctuate. However, peak date estimates appeared to stabilize, projecting a peak in mid-February, around early January 2016. This stable projection occurred before the true peak in the first week of March 2016, and it remained stable (±1 week) throughout the subsequent study period. By contrast, jurisdiction-specific peak estimates for Alberta (Supplementary Figure S8) and Ontario (Supplementary Figure S10) did not stabilize until after peaks had occurred, with the post-peak estimate for Ontario being 1 month earlier than the true peak.

Estimates of Control With a Fixed Reproduction Number

Estimates of d were obtained using a fixed estimate of R_0 (1.408) derived using all available data (Supplementary Figure 6). Using a fixed estimate of R_0, with the model used only to derive estimates of d, still produced parameter estimates that stabilized

Table 1. Estimates of R_0 and d for Populations and Generation Time Estimates Used in Analyses

Population and Generation Time	R0	d
Alberta cases only, 3-day generation time	1.372 (1.345–1.402)	0.0048 (0.0043–0.0052)
Ontario cases only, 3-day generation time	1.372 (1.347–1.399)	0.0047 (0.0043–0.0051)
Nova Scotia cases only, 3-day generation time	1.229 (1.184–1.278)	0.0045 (0.0033–0.00456)
Ottawa cases only, 3-day generation time	1.162 (1.139–1.187)	0.0022 (0.0018–0.0026)
Total population (Alberta, Ontario, and Nova Scotia cases), 3-day generation time	1.402 (1.378–1.428)	0.0047 (0.0044–0.0050)
Total population, 2-day generation time	1.244 (1.235–1.253)	0.0020 (0.0020–0.0022)
Total population, 4-day generation time	1.594 (1.564–1.625)	0.0085 (0.0080–0.0090)

Note: Estimates generated using complete seasonal data for each subgroup/assumption.
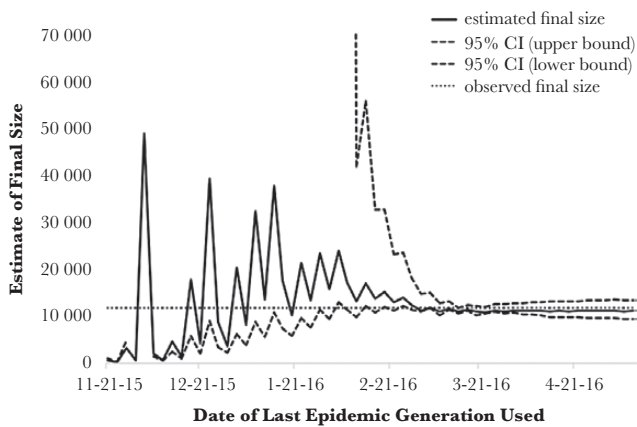
Figure 3. Final size projections for 2015–2016 influenza season. Projections of final epidemic size generated using the incidence decay with exponential adjustment (IDEA) model for summed influenza counts from Ontario, Alberta, and Nova Scotia (y-axis). The dates plotted on the x-axis are the dates of most recent influenza data used for model fitting. The solid black line represents projected size; dashed black lines represent upper- and lower-bound 95% confidence intervals (CIs). Upper-bound CI for final size before February 2016 are >100 000 and not shown on the graph. Dashed gray line represents true final size (11 686 cases).

near the end of January, providing no clear advantage over fitting both R_0 and d simultaneously.

DISCUSSION

The costs and health impact of seasonal influenza epidemics are substantial. Seasonal influenza epidemics vary in timing and severity from year to year, and it would be desirable to develop influenza forecasting methods that are accurate, reliable, and relatively simple and user friendly. Approaches to forecasting influenza have included (1) statistical models of varying complexity as
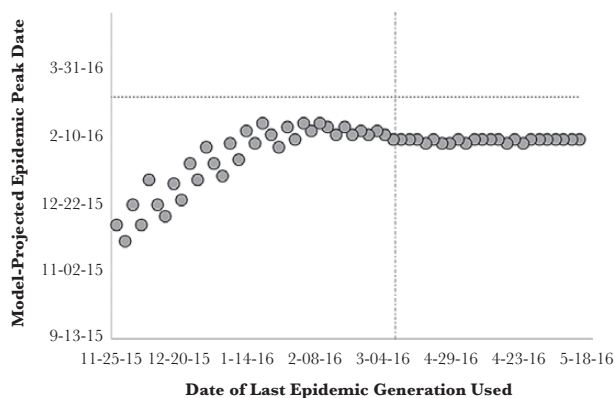

Figure 4. Peak date projections for 2015–2016 influenza season. Projections of final epidemic peak date generated using the incidence decay with exponential adjustment (IDEA) model for summed influenza counts from Ontario, Alberta, and Nova Scotia. The dates plotted on the x-axis are the dates of most recent influenza data used for model fitting; the y-axis represents the projected peak date. As noted in the text, peak date projections stabilized in mid-February before the true peak (first week of March 2016). Thus, a stable peak date was forecast before the true peak, but this forecast was persistently earlier than the true peak date.

well as (2) disease dynamic transmission models [17]. The performance of these approaches has been variable, and in a recent influenza forecasting exercise (Predict the Influenza Season Challenge) conducted under the auspices of the US Centers for Disease Control and Prevention, only 3 of 13 models correctly forecast influenza season onset, 1 of 13 correctly forecasts the season's peak, and 3 of 13 forecast epidemic size to within 1% [17].

In this study, we present a simple, phenomenological model for epidemic growth that we have described previously [10], but which was not included in the above-mentioned competition. The model we present above has the advantage of being sufficiently simple that it can be constructed in standard statistical spreadsheet software.

Using weekly pooled virological data from 3 Canadian regions and 1 city as inputs, we found that the model generated timely and fairly accurate estimates of influenza season size and duration. The parameters generated by the model seemed reasonable; our R_0 estimate for seasonal influenza (1.41) is consistent with estimates generated in other studies using more complex modeling approaches [17]. Although peak forecasts were off by approximately 2 weeks, forecasts had stabilized before the peak occurred, suggesting that accuracy may be improved with relatively simple adjustments in future; it may be that the persistent early peak projections generated by IDEA reflect the fact that the IDEA curve is symmetrical, whereas influenza epidemic curves may (as in this case) drop off sharply after the peak is reached. The IDEA model performed reasonably well in projecting final outbreak size. By early March 2016, we were able to estimate a final size of 11 002, compared with a total of 11 686 cases by the end of the outbreak, corresponding to error of approximately 6%, with final size estimates again having stabilized before the epidemic peak. However, confidence intervals around this estimate were wide with the upper bound confidence limit not dropping below 15 000 cases until mid-March (post-peak).

In jurisdictional analyses, we found what appeared to be a threshold effect for case counts in generating useful parameter estimates. Although estimates from our 2 largest jurisdictions (Ontario and Alberta) were fairly consistent with one another, and also with pooled overall estimates, estimates generated using data from smaller jurisdictions and/or those with more restrictive testing algorithms (Nova Scotia virology laboratory and Ottawa Public Health) was more challenging. The small number of cases reported from these areas was associated with low estimates of R_0. In general (and unsurprisingly), the smaller the sample size of the jurisdiction under analysis, the less stable, reliable, and timely projections appeared to be. Because frequency of testing or variability in testing practices may contribute to small test-positive numbers, it is unfortunate that we lacked testing data from most jurisdictions, which might have permitted us to evaluate the impact of tests per capita on usefulness of this modeling approach. Such explorations will be the subject of future work.

In addition, and as noted by an anonymous reviewer of an earlier draft of this paper, our early overestimation of reproduction numbers (and consequently, overestimation of d) early in the influenza season, is consistent with the tendency towards early overestimation of reproduction numbers noted by Mercer et al [18]. One possible reason for this overestimation relates to our "seeding" our models by assuming a single case at "time zero". Given the well described underdiagnosis of influenza [19], this heuristic is may be problematic, and if our outbreak is seeded by >1 initial case that would result in reduction in early estimates of R_0. It is unclear whether there may be some empirically derived inflation factor that might allow improved early performance (and timelier characterization of epidemics) by the IDEA model, but this represents an interesting question for future research.

CONCLUSIONS

The performance of our model during this exercise, while somewhat encouraging, may not predict performance in future influenza seasons. The 2015–2016 influenza season was atypical relative to other recent influenza seasons, with a late start and relatively low impact [20]. The mechanisms underlying these atypical dynamics are not clear, but better understanding of the impact of between-season variation in predominating circulating strain, and vaccine match and uptake, may in turn permit more accurate and timely model projections. This project focused on the ability to forecast influenza epidemics in real time, but better understanding of factors influencing model performance will likely also need to include historical data from well characterized influenza seasons, and we are currently engaged in such work.

References

1. Clark NM, Lynch JP 3rd. Influenza: epidemiology, clinical features, therapy, and prevention. Semin Respir Crit Care Med **2011**; 32:373–92.
2. Levenson JE, Kaul DR, Saint S, et al. Clinical problem-solving. A shocking development. N Engl J Med **2013**; 369:2253–8.
3. Simonsen L, Clarke MJ, Schonberger LB, et al. Pandemic versus epidemic influenza mortality: a pattern of changing age distribution. J Infect Dis **1998**; 178:53–60.
4. Molinari NA, Ortega-Sanchez IR, Messonnier ML, et al. The annual impact of seasonal influenza in the US: measuring disease burden and costs. Vaccine **2007**; 25:5086–96.
5. Simonsen L, Fukuda K, Schonberger LB, Cox NJ. The impact of influenza epidemics on hospitalizations. J Infect Dis **2000**; 181:831–7.
6. Fisman D. Modelling an influenza pandemic: A guide for the perplexed. CMAJ **2009**; 181:171–3.
7. Chowell G, Sattenspiel L, Bansal S, Viboud C. Mathematical models to characterize early epidemic growth: A review. Phys Life Rev **2016**; 18:66–97.
8. Majumder MS, Rivers C, Lofgren E, Fisman D. Estimation of MERS-coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: insights from publicly available data. PLoS Curr **2014**; 6.
9. Fisman D, Tuite A. Projected impact of vaccination timing and dose availability on the course of the 2014 West African ebola epidemic. PLoS Curr **2014**; 6.
10. Fisman DN, Hauck TS, Tuite AR, Greer AL. An IDEA for short term outbreak projection: nearcasting using the basic reproduction number. PLoS One **2013**; 8:e83622.
11. Tuite AR, Tien J, Eisenberg M, et al. Cholera epidemic in Haiti, 2010: using a transmission model to explain spatial spread of disease and identify optimal control interventions. Ann Intern Med **2011**; 154:593–601.
12. Bolker B. Bblme: Tools for General Maximum Likelihood Estimation. **2016**; Available at: https://cran. r-project. org/web/packages/bbmle/index.html.
13. Tuite AR, Fisman DN. The IDEA model: A single equation approach to the Ebola forecasting challenge. Epidemics **2016**, http://dx.doi.org/10.1016/j.epidem.2016.09.001.
14. Cowling BJ, Fang VJ, Riley S, et al. Estimation of the serial interval of influenza. Epidemiology **2009**; 20:344–7.
15. Centers for Disease Control and Prevention. Summary of the 2015–2016 Influenza Season. Available at: https://www.cdc.gov/flu/about/season/flu-season-2015-2016.htm. Accessed 17 January 2017.
16. Biggerstaff M, Cauchemez S, Reed C, et al. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infect Dis **2014**; 14:480.
17. Biggerstaff M, Alper D, Dredze M, et al. Results from the centers for disease control and prevention's predict the 2013–2014 Influenza Season Challenge. BMC Infect Dis **2016**; 16:357.
18. Mercer GN, Glass K, Becker NG. Effective reproduction numbers are commonly overestimated early in a disease outbreak. Stat Med **2011**; 30:984–94.
19. Penttinen P, Pebody R. Influenza case definitions - optimising sensitivity and specificity. Euro Surveill **2015**; 20:21148.
20. Centers for Disease Control and Prevention (CDC). Interventions to increase influenza vaccination of health-care workers--California and Minnesota. MMWR Morb Mortal Wkly Rep **2005**; 54:196–9.