ANNUAL REVIEWS

Influenza Virus: Tracking, Predicting, and Forecasting

Sheikh Taslim Ali and Benjamin J. Cowling

World Health Organization Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China; email: bcowling@hku.hk

Keywords

surveillance systems, data assimilation, transmission dynamics, prediction, forecasting, influenza

Abstract

Influenza is a common respiratory infection that causes considerable morbidity and mortality worldwide each year. In recent years, along with the improvement in computational resources, there have been a number of important developments in the science of influenza surveillance and forecasting. Influenza surveillance systems have been improved by synthesizing multiple sources of information. Influenza forecasting has developed into an active field, with annual challenges in the United States that have stimulated improved methodologies. Work continues on the optimal approaches to assimilating surveillance data and information on relevant driving factors to improve estimates of the current situation (nowcasting) and to forecast future dynamics.

1. INTRODUCTION

Seasonal influenza epidemics mainly caused by influenza A and B viruses result in ∼3–5 million cases of severe illness and 290,000–650,000 deaths worldwide each year (87). Seasonal influenza viruses circulate annually during the winter in temperate locations, but influenza seasonality is not well defined in tropical and subtropical locations and can circulate year-round (11, 78). In contrast with seasonal influenza, novel influenza A strains capable of sustained person-to-person transmission arise occasionally in global pandemics.

Influenza surveillance is an important public health activity. Influenza surveillance data can be used to signal the start and end of an influenza season, to describe the impact of influenza seasons, and to indicate the impact of control and mitigation measures. Laboratory-based influenza surveillance can also provide samples of circulating strains in different parts of the world, which is useful for making decisions on vaccine strain selection and for monitoring the prevalence of resistance to antiviral drugs. The World Health Organization (WHO) Global Influenza Surveillance Network (GISN) was founded in 1952 and renamed the Global Influenza Surveillance and Response System in 2011 to coordinate global surveillance activities (99).
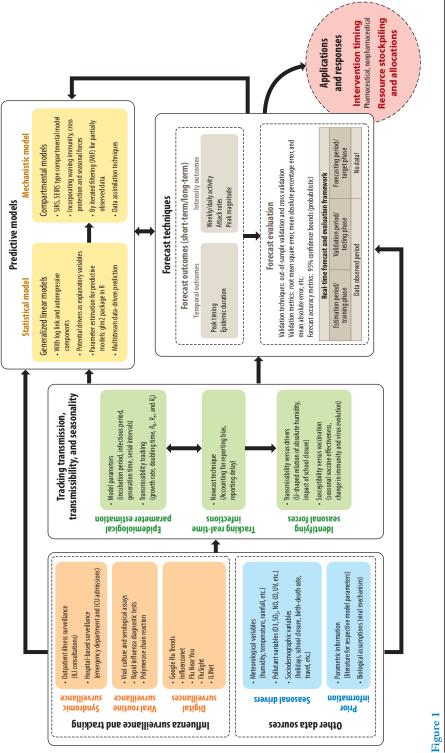
The science of surveillance has been rapidly evolving in recent years (40, 76). At the same time, improvements in computational skill allow advanced modeling techniques for prediction and forecasting by incorporating multiple streams of data and complex underlying dynamics. These improvements present new opportunities not only for understanding the current situation, referred to as nowcasting, but also for forecasting what might happen in future weeks, for example, determining when the influenza season will reach a peak. Providing reliable forecasts of future influenza activity can be extremely informative for allocating health resources and judging public health measures. **Figure 1** illustrates the different steps of tracking, predicting, and forecasting influenza activity. In this review, we attempt to address, at each step, the potential uses, challenges, and possible research gaps that need to be pursued in future courses of tracking, predicting, and forecasting influenza virus circulation in the community.
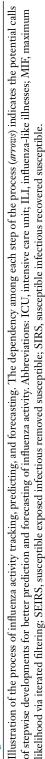
2. INFLUENZA SURVEILLANCE AND TRACKING

The WHO defines public health surveillance as "an ongoing, systematic collection, analysis and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice and is undertaken to inform disease prevention and control measures" (89). Public health surveillance has evolved considerably in recent decades (23).

Influenza surveillance is a common public health activity done around the world. The WHO coordinates global influenza surveillance and provides recommendations for the conduct of influenza surveillance in Member States (86). Having learned from the 1918–1919 pandemic, the WHO established the Global Influenza Programme in 1947 and the Global Influenza Surveillance Network (GISN) in 1952. The latter was renamed as the Global Influenza Surveillance and Response System (GISRS) in 2011 and has expanded to include 143 institutions in 113 Member States (88).

Broadly speaking, influenza surveillance is carried out via two main approaches. First, to understand the occurrence of influenza illnesses in the community, syndromic surveillance systems collect information from outpatient clinics or hospitals and describe patterns in incidence over time. Second, laboratory surveillance systems test samples collected from influenza patients and provide information on the virologic characteristics of circulating strains. This information can be used in vaccine strain selection and to monitor resistance against antiviral drugs. Some countries have established multiple sources of influenza surveillance data, which can be synthesized to provide more detailed information on influenza dynamics, improving the value of this information (**Figure 1**).

Figure 1

Illustration of the process of influenza activity tracking, predicting, and forecasting. The dependency among each step of the process (*arrows*) indicates the potential calls of stepwise developments for better prediction and forecasting of influenza activity. Abbreviations: ICU, intensive care unit; ILI, influenza-like illnesses; MIF, maximum likelihood via iterated filtering; SEIRS, susceptible exposed infectious removed susceptible; SIRS, susceptible infectious recovered susceptible.

2.1. Syndromic Surveillance

The classic approach to syndromic influenza surveillance is sentinel outpatient surveillance. In this approach, a network of outpatient clinics ("sentinels") provide information at a specified frequency, typically weekly, on the number of patients they have attended with influenza-like illnesses (ILI) (6, 31, 61, 67, 75). A typical definition of an ILI is recent-onset fever plus cough. Examples of this approach include the US Outpatient Influenza-like Illness Surveillance Network (ILINet) (20) and the European Surveillance System (TESSy) at the European Centre for Disease Prevention and Control (ECDC) (31). The catchment populations of the sentinel clinics are the preferred denominator for these data so that the illness rates provide information on incidence rates of medically attended influenza in the population. An alternative denominator in locations with unclear catchment populations would be total consultations. In either case, rates of ILI in the community indicate the start and end of influenza seasons, and comparisons between years can indicate which seasons are more intense than others. As an alternative approach to tracking levels of influenza in the underlying community, similar data can be captured from hospitals that have admitted patients with influenza (40, 79). More recently, other approaches to syndromic surveillance include processing electronic medical records or telenursing calls for relevant diagnostic codes (15, 20, 39, 46, 51, 53, 98).

2.2. Laboratory Surveillance

Laboratory-based surveillance is an essential component of many influenza surveillance systems. In the 1940s, growing virus in culture allowed antigenic characterization of strains and supported vaccine development. This method is still widely used today as part of the process for antigenic and genetic characterization of influenza viruses (21). In 1983, the development of the polymerase chain reaction (PCR) technique by Kary Mullis revolutionized the field of infectious disease diagnosis. This molecular technique detects virus-specific RNA sequences rather than viral antigens or antibodies and can provide results within a few hours. Most laboratory surveillance of influenza is now done by PCR.

One other approach bears mention, and that is serological analysis. A few weeks after a person recovers from an influenza virus infection, they will develop an antibody response against the infecting strain, which can be detected in their peripheral blood. Testing of blood samples is a diagnostic tool but, more importantly, may be used to provide surveillance data for a population. For example, periodic collection of blood samples can allow researchers to analyze levels of infection or immunity in the population and indicate whether there might be vulnerability to certain strains that are prevalent in other parts of the world (85).

2.3. Digital Surveillance and Emerging Data Sources

These conventional influenza surveillance approaches primarily describe outbreaks of influenza based on reports from doctors or clinicians. More recently, researchers have explored the potential for electronic data sources to provide more timely information on levels of influenza circulation in the community. Examples of this approach include trends derived from (*a*) searches (e.g., Google Search Trends, Wikipedia page views), (*b*) social media postings (e.g., Facebook posts, tweets), and (*c*) participatory surveillance efforts (e.g., Flu Near You, Influenzanet) (2). These types of surveillance enlist ProMED-mail (50), De Grote Griepmeting (45), webpage views used for influenza surveillance (42), Google Ads click rate used for surveillance (32), HealthMap (14), Google Flu Trends (25, 57), Flu Near You (77), FluSight (19), etc. The prime objective of these systems was to provide timely reports of influenza incidence in local areas (16, 25, 77), and the resulting data were often used for influenza prediction and forecasting.

Digital surveillance data do have limitations, though. One of the problems occurs when these systems capture changes in public awareness of the disease rather than changes in incidence of the disease per se. For example, the inaccurate predictions by Google Flu Trends led to its discontinuation in 2015 (35) after it failed to detect the influenza A (H1N1) pandemic in 2009 and greatly overestimated the peak intensity of the 2012/2013 season (57). Moreover, data from digital surveillance often suffer from problems with stability, dependency, confounding in search terms, and representativeness (13, 36, 64). Despite these challenges, digital surveillance data are comparatively lower in cost and could supplement other more traditional sources of information rather than replace them. One growing research area involves the development of hybrid systems, which couple traditional surveillance data with data from digital surveillance tools, including search queries, social media posts, and crowdsourcing (2).

2.4. Multistream Data Assimilation and Synthesis

While ILI rates provide information on levels of influenza activity in the community, these rates rarely decline completely to zero outside of the influenza season because of the circulation of other respiratory virus infections. To provide a metric that better correlates with the incidence of influenza virus infections in the community, Goldstein et al. (37) introduced the ILI+ proxy by multiplying ILI rates with rates of influenza detection in the laboratory. This ILI+ proxy was found to be a closer linear correlate of influenza virus infections in the 2009 influenza pandemic than were ILI rates or laboratory detection rates alone (91), and it has now been widely adopted as a good measure of representation for the influenza epidemic curve in a population (3, 4, 44, 66, 72, 92, 94, 96, 97).

While the ILI+ proxy can provide a reasonable representation of the shape of the epidemic curve for an influenza epidemic, the area under this curve would differ from the infection attack rate in the population by an unknown constant of proportionality, sometimes referred to as the reporting rate. The absolute rate of influenza virus infections in a population is challenging to estimate from these data streams because of (*a*) the uncertainties about the age- and strain-specific likelihoods of consulting a physician upon contracting an influenza virus infection, (*b*) the sensitivity of the laboratory testing of respiratory specimens, and (*c*) the lack of patient age recording for influenza-positive specimens in many countries, including the United States (60). One possible solution would be to integrate information from serological surveillance with data on ILI consultations and virological testing of respiratory specimens to allow complete estimation of influenza infection rates (both symptomatic and asymptomatic) during epidemics (7, 17, 58).

More timely information on influenza activity via digital surveillance has the potential to allow a more accurate measure of the incidence of influenza virus infections in the community when coupled with data from traditional surveillance. This data assimilation technique opens up an opportunity to develop the hybrid digital surveillance systems (47, 74) using web-based surveillance with more traditional sources. Santillana et al. (68) combined Google, Twitter, and Flu Near You data with ILI rates from symptomatic surveillance and found an improved measure of influenza activity compared with using Google Flu Trends data alone. Several studies addressed the development of the hybrid digital surveillance system by accounting for methodological biases, including stability, dependency (13), confounding in search terms (69), and sampling bias or representativeness, along with the above data integration techniques. In comparison, big-data, cloud-based electronic health record data sources are often available in almost real time, can be integrated with ILI estimates of search engine data, and can provide more accurate predictions of influenza intensity up to four weeks in advance (95). Data on the potential drivers of influenza transmission, such as data on meteorological factors, could be incorporated with these surveillance data to improve

predictions and forecasting (5, 43, 71, 72). We further discuss these seasonal drivers and their roles in predicting influenza in Section 3.2 and Section 4.

3. INFLUENZA TRANSMISSION DYNAMICS: TRANSMISSIBILITY AND SEASONALITY

Along with good-quality influenza surveillance data, having a clear understanding of the transmission dynamics of influenza, and how these dynamics might be influenced by social mixing patterns or extrinsic factors that drive seasonality, can also be valuable.

3.1. Tracking Real-Time Transmissibility

Influenza is an acute contagious infection of the airways and lungs caused by an influenza virus infection. Influenza transmission dynamics are characterized by several epidemiological parameters estimated from available surveillance data. These dynamics include the incubation period, infectious period, serial interval, generation time, exponential growth rate, and reproduction number, among others. Reliable information on these parameters is important for influenza prediction and forecasting studies, as discussed below.

The most widely used measure of transmissibility for any infectious disease is the reproduction number (R), which is defined as the average number of secondary cases generated per typical infectious case. In theory, R depends on the population susceptibility (or immunity). When R is estimated for a population that is entirely susceptible to infection (no prior immunity and no intervention implemented), it is referred to as the basic reproduction number (R_0). In contrast, the effective reproduction number (R_e) is estimated in a population with underlying immunity and accounts for reduced population susceptibility to infection (12). To quantify real-time transmissibility, the time-varying instantaneous reproduction number (R_t) can be estimated as the effective reproductive number on day t (26, 55, 80). Biggerstaff et al. (9) reviewed estimates of R for seasonal, pandemic, and zoonotic influenza and presented the variation in the estimates by the assumption on generation times and case definitions, along with the case data used from different locations.

3.2. Seasonality

Influenza seasonality describes a periodic surge in influenza virus infections during certain times of the year (28). In temperate regions, influenza virus infections are characterized by a regular influenza season (lasting for a total of 5–10 weeks with a single peak), typically in the winter months during which infections increase significantly while remaining at very low levels throughout the rest of the year. In tropical regions, on the other hand, influenza epidemics are less regular, and the timing and duration are not well defined. Some locations can experience year-round circulation or multiple epidemics in a single year (22, 78, 90).

To explain these spatiotemporal differences in the seasonality of influenza virus circulation, several intrinsic and extrinsic drivers have been proposed. Intrinsic biological drivers of seasonality include viral evolution and the immune response to the virus and seasonal host health (48). Slow but steady antigenic changes in circulating influenza viruses, referred to as antigenic drift, occur in response to selection pressure from population immunity (27). Limited immune cross-protection against antigenic variants can generate cyclical patterns, closely resembling the seasonal patterns of influenza virus circulation (33). Seasonal variations in host health and physiological status are quite common; even the immune system may experience a pattern of ebb and flow that could leave a host more vulnerable to infection (34, 41, 48). By comparison, extrinsic drivers include

meteorological, environmental (e.g., pollutants), and social determinants (e.g., school closures) as well as vaccination programs. These extrinsic drivers could affect the survivability and viability of the virus in different ambient settings or could change the potential for infection to be transmitted from one individual to another. These external seasonal determinants can contribute to the seasonality of influenza virus transmission and would be important complementary data sources for influenza prediction and forecasting (72, 73, 78), discussed further in Section 4 below.
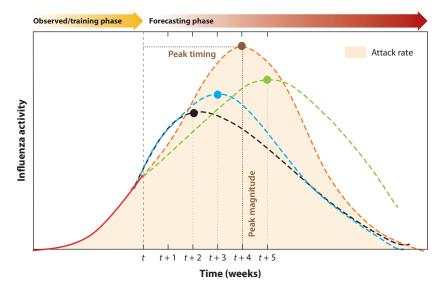
Although the seasonal recurrence of influenza epidemics in temperate regions has been comparatively well characterized for decades, explanations of seasonal influenza virus circulation in the tropics and subtropics are still poorly understood (30). For example, low absolute humidity in temperate locations is associated with influenza circulation; however, in the tropics, influenza epidemics occur during periods with both low humidity and very high humidity, resulting in a U-shaped association between humidity and influenza (78).

4. INFLUENZA PREDICTION AND FORECASTING

Data on influenza surveillance and knowledge on influenza transmission dynamics can be used to construct models for predicting future influenza activity or for creating counterfactual scenarios about what might have happened in alternative circumstances, for example, with or without the implementation of particular public health measures. The prediction of influenza activity involves the art of integrating these data and information under model development exercises. In general, studies on the prediction of influenza are classified into two categories: projection and forecasting (52). A projection is a realization or comparison of what would happen under certain assumptions and hypotheses (e.g., effectiveness of control measures), whereas a forecast is a quantitative estimate of what will happen in the future (e.g., upcoming influenza activity). The studies on influenza projection have a longer history, whereas influenza forecasting is comparatively newer. In the past 10 years, there has been a growing effort to develop the systems and methods for forecasting different characteristics of influenza epidemics. The ability to predict future incidence, timing of epidemics, and peaks of influenza epidemics in a timely manner could be extremely valuable for health care planning and resource allocation.

4.1. Development of Predictive Models for Influenza Forecasting

Influenza forecasting, as illustrated in **Figure 1**, requires predictive models along with influenza surveillance data. The advances in computational facilities and data assimilation techniques allow complex algorithms and models to be constructed and applied. Models are usually trained on data from a training period and used for predicting future events (forecast outcomes) during a forecasting period. The most accurate forecasts would generally, although perhaps not always, be made by a model that has good predictive performance. The forecast outputs are of two types: temporal outcomes (e.g., peak timing, epidemic duration) and intensity outcomes (e.g., peak magnitude, attack rates). One early study was reported by Longini et al. (49), who evaluated ILI activity and the peak timing of the 1968/1969 Hong Kong influenza pandemic. Before the 2009 influenza pandemic, several notable works were published on forecasting influenza activity (1, 62, 83), peak timing (1, 38, 81), and peak magnitude and epidemic duration (1, 38). The number of studies increased significantly after the 2009 influenza pandemic, highlighting the increased recognition and importance given to preparedness and public health planning strategies for influenza. The models in these studies have their own advantages and limitations, which are reviewed extensively by Nsoesie et al. (56) and Chretien et al. (24), who also highlighted the prospective opportunities in the area of influenza forecasting.

Figure 2

Illustration of forecasting outcomes and their measures to quantify the forecasting performance under possible counterfactual scenarios. The dashed curves in different colors are the forecast of influenza activities by four different models. Presented are the results of four models on forecasting the peak timing (*bold dots*), peak magnitude (*vertical dotted lines*), and attack rate (*area under the curves*) of a typical outbreak. For a particular model, the forecast outcomes are presented as peak timing from the observed phase (*brown horizontal dotted line*), peak magnitude (*brown vertical line*), and attack rate (*light orange*). Abbreviation: *t*, time.

In 2013, the US Centers for Disease Control and Prevention (CDC) launched an annual national-level influenza forecasting competition, Predict the Influenza Season Challenge, as a collaboration with academic researchers (18). The CDC set forecasting targets relevant to public health decision makers through real-time forecast (1–4 weeks ahead) of the influenza seasons in the United States; the forecast outcomes include the timing of the onset of the influenza season, the timing of the peak, and the peak intensity (**Figure 2**). More than 30 teams currently submit their models for competition each year with the objectives of continuing to improve the research gaps on forecasting model development, evaluation, and adoption by decision makers, including the need to develop standardized metrics to assess forecast accuracy and consistent ways to communicate forecasts and their uncertainty. One particular emphasis of that consortium was on the probabilistic forecasting framework (i.e., what is the probability that the influenza season will peak next week, or in 2 weeks, etc.?) rather than simple all-or-nothing forecasts (8, 10). Recently, Xu et al. (93) investigated predictive utility of online social media and web search queries to forecast new cases of ILI in Hong Kong. Such models with alternative or parallel sources of information (such as search engine queries) could improve real-time influenza prediction and could be deployed for forecasting (54). A rigorous computational framework for real-time forecasts of seasonal influenza outbreaks, using a data assimilation technique (either the ensemble adjustment Kalman filter or modified particle filtering under a compartmental model), commonly applied in numerical weather prediction, was developed by Shaman & Karspeck (71) and further developed using other variables (e.g., climatic, seasonal forces) (96). As the disease dynamics are not often completely observed, researchers have proposed an alternative approach to considering influenza

transmission as a partially observed process through a state-space model motivated by a compartmental model with a Bayesian framework for inference and forecasting (59). Incorporating evolutionary change into such a mechanistic epidemiological model setup improved the forecasting accuracy of influenza A (H3N2) epidemics in the United States (29).

Many years of data are available from the CDC forecasting challenge, allowing some general conclusions on the performance of different types of models (8, 10). In earlier years, the model performances for this forecasting challenge were evaluated relative to the results of the ILINet benchmark. The performance of the models submitted for the 2014/2015 influenza season presented a wide variation in accuracy; the most accurate forecast models (mechanistic and statistical) used data on ILINet, specific humidity, and crowdsourced forecasts in US regions. For the most recent influenza seasons, the submitted models provided more accurate predictions than those using ILINet's historical baseline (70); therefore, the CDC proposed to implement those top-performing models for the FluSight system. Probabilistic forecast distributions allow for a quantitative evaluation of accuracy, which can be used to compare and communicate forecast performance (10).

Broadly speaking, two general approaches have been used to make forecasts in these studies. One approach uses statistical and phenomenological methods, such as time series models, generalized linear models, classification algorithms, regression models, and Bayesian networks. The second approach uses mechanistic models such as compartmental models and agent-based models. While mechanistic approaches can sometimes better account for nonlinear infection dynamics, they also require a much larger number of parameters and assumptions.

Among the statistical approaches, predictive models are typically developed where the response variable is a relevant measure of influenza activity (e.g., ILI+ proxy). This response variable is linked by some type of regression model to a number of covariates, possibly with nonlinear components, to allow for infection dynamics. The forecast validations are generally performed with out-of-sample validation, where the model parameters are recomputed for each forecast by using only the training data from the previous weeks and the upcoming weeks for validation. Forecast accuracy is evaluated by considering one or more of the following measures, including root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE).

In a typical mechanistic approach, influenza activity has to somehow be translated to a measure of incidence, with additional assumptions about population immunity. Transmissibility can then be estimated in the constructed model, accounting for covariates. The probabilistic forecast distributions allowed for a quantitative evaluation of accuracy, which can generally be used to compare and communicate forecast performance. These models have their own advantages and challenges. For example, in the first and second years of the CDC challenges, ~56% of proposed models were statistical models and 44% were mechanistic models. Forecast accuracy of these models varied widely even evaluated for the same influenza season. Among the proposed models, a statistical model and a mechanistic model were found to have the most accurate forecasts for peak timing and seasonal targets (8, 10). The statistical model used ILINet and crowdsourced data to produce many different influenza forecasts and finally to generate an aggregate forecast. The mechanistic approach used a compartmental model with ILINet and specific humidity data, combined with three different ensemble filter algorithms (12 model–filter combinations).

While disease forecasting has developed as an innovative discipline over the last decade, it is not yet fully integrated into public policy. Forecasting models for influenza continue to improve, and policy makers should have growing confidence in their ability to make reliable short-term forecasts based on these models.

4.2. The Future of Prediction and Forecasting

Better forecasting performance depends on a series of steps (**Figure 1**), starting from (*a*) real-time detection and tracking of influenza cases; (*b*) understanding of the transmission dynamics for estimating model parameters outside/inside predictive models; (*c*) identification of the intrinsic and extrinsic drivers (predictors) of influenza transmissibility and seasonality; and, finally, (*d*) application of predictive models, which requires integrating all this information simultaneously. Therefore, the forecasting model development exercise should incorporate parameter optimization and reparameterization, evaluation of comparative models using common input data with validation and accuracy metrics, calibration of analogous monitoring frameworks for implementing a probabilistic forecast, and integration of multistream data in the model under different forecasting schemes. In fact, the CDC influenza forecasting challenge, which is focused on short-term (up to four weeks) prediction of the trajectory of influenza outbreaks in US regions, encouraged the use of social media data to predict influenza (8, 10, 19). In 2019, the study by Reich et al. (65) is one such example of multiyear (seven years) forecasting of seasonal influenza, which featured a variety of modeling approaches with consistent model formulations and forecasting targets throughout the study period. Their modeling framework includes a portfolio of both statistical and mechanistic models and a range of data streams, including digital surveillance, meteorological data, and social media content such as Google queries and Twitter posts, along with case counts. Therefore, the study by Reich et al. (65) has set a benchmark for future forecasting model exercises, which will be much-needed for further improving the models (84).

Forecasting performance of predictive models can vary by location (65). It would be informative to disentangle the reasons for these regional differences by exploring reporting artifacts or heterogeneities in transmission dynamics along with the various possible drivers, including demographic and environmental differences and social mixing patterns in these regions. Because no two successive influenza seasons are identical, it can be important to include in the models some key biological and epidemiological features, such as population immunity, antigenic changes in the virus, and vaccine coverage. Improved prediction approaches could even open up the possibility of forecasting influenza dynamics in the tropics. In general, statistical models seem to have better performance for short-term forecasting, whereas mechanistic models can be more flexible and tend to allow more accurate longer-term forecasts (65, 84). So far, prediction models have been based mostly on ILI or ILI+ proxies of influenza activity (72). Prediction accuracy can vary depending on which proxy measure is used, and neither of them may be a gold standard reflection of the underlying incidence of influenza virus infections in the community. These proxies can suffer reporting biases, biases due to health care–seeking behavior, missing data, and the contribution of other respiratory virus infections (84). Another general issue is the typical 1–2-week lag in influenza surveillance data, such that forecasting models are even required to estimate levels of influenza activity today (known as nowcasting) (63, 82). There remain many opportunities to improve influenza prediction and forecasting models.

5. CONCLUSIONS

In this review, we have outlined the characteristics of influenza surveillance systems and the approaches that have been used to predict and forecast influenza activity. Each step of the surveillance, analysis, and interpretation process is important for influenza prediction and forecasting (**Figure 1**). Influenza prediction and forecasting challenges have been established and run mostly by public health institutions; however, in the future, we may see broader collaborative networks emerge, including infectious disease modelers as well as the end users in local public health agencies and health care providers.

DISCLOSURE STATEMENT

ACKNOWLEDGMENTS

LITERATURE CITED

1. Aguirre A, Gonzalez E. 1992. The feasibility of forecasting influenza epidemics in Cuba. *Mem. Inst. Oswaldo Cruz* 87:429–32
2. Aiello AE, Renson A, Zivich PN. 2020. Social media- and Internet-based disease surveillance for public health. *Annu. Rev. Public Health* 41:101–18
3. Ali ST, Cowling BJ, Lau EHY, Fang VJ, Leung GM. 2018. Mitigation of influenza B epidemic with school closures, Hong Kong, 2018. *Emerg. Infect. Dis.* 24:2071–73
4. Ali ST, Wu P, Cauchemez S, He D, Fang VJ, et al. 2018. Ambient ozone and influenza transmissibility in Hong Kong. *Eur. Respir. J.* 51:1800369
5. Axelsen JB, Yaari R, Grenfell BT, Stone L. 2014. Multiannual forecasting of seasonal influenza dynamics reveals climatic and evolutionary drivers. *PNAS* 111:9538–42
6. Babakazo P, Kabamba-Tshilobo J, Wemakoy EO, Lubula L, Manya LK, et al. 2019. Evaluation of the influenza sentinel surveillance system in the Democratic Republic of Congo, 2012–2015. *BMC Public Health* 19:1652
7. Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. 2013. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modelling study. *PLOS Med.* 10:e1001527
8. Biggerstaff M, Alper D, Dredze M, Fox S, Fung IC, et al. 2016. Results from the Centers for Disease Control and Prevention's Predict the 2013–2014 Influenza Season Challenge. *BMC Infect. Dis.* 16:357
9. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. 2014. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Infect. Dis.* 14:480
10. Biggerstaff M, Johansson M, Alper D, Brooks LC, Chakraborty P, et al. 2018. Results from the second year of a collaborative effort to forecast influenza seasons in the United States. *Epidemics* 24:26–33
11. Bloom-Feshbach K, Alonso WJ, Charu V, Tamerius J, Simonsen L, et al. 2013. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. *PLOS ONE* 8:e54445
12. Boëlle PY, Ansart S, Cori A, Valleron A-J. 2011. Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: a review. *Influenza Other Respir. Viruses* 5:306–16
13. Broniatowski DA, Paul MJ, Dredze M. 2013. National and local influenza surveillance through Twitter: an analysis of the 2012–2013 influenza epidemic. *PLOS ONE* 8:e83672
14. Brownstein JS, Freifeld CC, Reis BY, Mandl KD. 2008. Surveillance sans frontières: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLOS Med.* 5:e151
15. Brownstein JS, Kleinman KP, Mandl KD. 2005. Identifying pediatric age groups for influenza vaccination using a real-time regional surveillance system. *Am. J. Epidemiol.* 162:686–93
16. Carneiro HA, Mylonakis E. 2009. Google Trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin. Infect. Dis.* 49:1557–64
17. Cauchemez S, Valleron A-J, Boëlle P-Y, Flahault A, Ferguson NM. 2008. Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* 452:750–54

18. CDC (Cent. Dis. Control Prev.). 2013. About CDC's flu forecasting efforts. *Influenza (Flu)*. **https://www.cdc.gov/flu/weekly/flusight/about-flu-forecasting.htm**

19. CDC (Cent. Dis. Control Prev.). 2013. *CDC competition encourages use of social media to predict flu*. Press Release, Nov. 25, CDC, Atlanta. **https://www.cdc.gov/flu/news/predict-flu-challenge.htm**

20. CDC (Cent. Dis. Control Prev.). 2020. Flu activity & surveillance. *Influenza (Flu)*. **https://www.cdc.gov/flu/weekly/fluactivitysurv.htm**

21. CDC (Cent. Dis. Control Prev.). 2020. Overview of influenza testing methods. *Influenza (Flu)*. **https://www.cdc.gov/flu/professionals/diagnosis/overview-testing-methods.htm#culture**

22. Chew FT, Doraisingham S, Ling AE, Kumarasinghe G, Lee BW. 1998. Seasonal trends of viral respiratory tract infections in the tropics. *Epidemiol. Infect.* 121:121–28

23. Choi BCK. 2012. The past, present, and future of public health surveillance. *Scientifica (Cairo)* 2012:875253

24. Chretien J-P, George D, Shaman J, Chitale RA, McKenzie FE. 2014. Influenza forecasting in human populations: a scoping review. *PLOS ONE* 9:e94130

25. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. 2011. Assessing Google Flu Trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PLOS ONE* 6:e23610

26. Cori A, Ferguson NM, Fraser C, Cauchemez S. 2013. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* 178:1505–12

27. Cox NJ, Subbarao K. 1999. Influenza. *Lancet* 354:1277–82

28. Dowell SF, Ho MS. 2004. Seasonality of infectious diseases and severe acute respiratory syndrome—what we don't know can hurt us. *Lancet Infect. Dis.* 4:704–8

29. Du X, King AA, Woods RJ, Pascual M. 2017. Evolution-informed forecasting of seasonal influenza A (H3N2). *Sci. Transl. Med.* 9:eaan5325

30. Dushoff J, Plotkin JB, Levin SA, Earn DJD. 2004. Dynamical resonance can account for seasonality of influenza epidemics. *PNAS* 101:16915–16

31. ECDPC (Eur. Cent. Dis. Prev. Control). 2020. Sentinel surveillance. *Seasonal Influenza*. **https://www.ecdc.europa.eu/en/seasonal-influenza/surveillance-and-disease-data/facts-sentinel-surveillance**

32. Eysenbach G. 2006. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu. Symp. Proc.* 2006:244–48

33. Ferguson NM, Galvani AP, Bush RM. 2003. Ecological and immunological determinants of influenza evolution. *Nature* 422:428–33

34. Fitch WM, Leiter JM, Li XQ, Palese P. 1991. Positive Darwinian evolution in human influenza A viruses. *PNAS* 88:4270–74

35. Flu Trends Team. 2015. The next chapter for flu trends. *Google AI Blog*, Aug. 20. **https://ai.googleblog.com/2015/08/the-next-chapter-for-flu-trends.html**

36. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457:1012–14

37. Goldstein E, Cobey S, Takahashi S, Miller JC, Lipsitch M. 2011. Predicting the epidemic sizes of influenza A/H1N1, A/H3N2, and B: a statistical method. *PLOS Med.* 8:e1001051

38. Hall IM, Gani R, Hughes HE, Leach S. 2007. Real-time epidemic forecasting for pandemic influenza. *Epidemiol. Infect.* 135:372–85

39. Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, Weiss D. 2004. Syndromic surveillance in public health practice, New York City. *Emerg. Infect. Dis.* 10:858–64

40. Hiller KM, Stoneking L, Min A, Rhodes SM. 2013. Syndromic surveillance for influenza in the emergency department—a systematic review. *PLOS ONE* 8:e73832

41. Ina Y, Gojobori T. 1994. Statistical analysis of nucleotide sequences of the hemagglutinin gene of human influenza A viruses. *PNAS* 91:8388–92

42. Johnson HA, Wagner MM, Hogan WR, Chapman W, Olszewski RT, et al. 2004. Analysis of Web access logs for surveillance of influenza. *Stud. Health Technol. Inform.* 107:1202–6

43. Kandula S, Pei S, Shaman J. 2019. Improved forecasts of influenza-associated hospitalization rates with Google Search Trends. *J. R. Soc. Interface* 16:20190080

44. Kandula S, Yamana T, Pei S, Yang W, Morita H, Shaman J. 2018. Evaluation of mechanistic and statistical methods in forecasting influenza-like illness. *J. R. Soc. Interface* 15:20180174

45. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, et al. 2017. Influenzanet: citizens among 10 countries collaborating to monitor influenza in Europe. *JMIR Public Health Surveill.* 3:e66

46. Lazarus R, Kleinman K, Dashevsky I, Adams C, Kludt P, et al. 2002. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emerg. Infect. Dis.* 8:753–60

47. Lazer D, Kennedy R, King G, Vespignani A. 2014. The parable of Google Flu: traps in big data analysis. *Science* 343:1203–5

48. Lofgren E, Fefferman NH, Naumov YN, Gorski J, Naumova EN. 2007. Influenza seasonality: underlying causes and modeling theories. *J. Virol.* 81:5429–36

49. Longini IM Jr., Fine PE, Thacker SB. 1986. Predicting the global spread of new infectious agents. *Am. J. Epidemiol.* 123:383–91

50. Madoff LC, Woodall JP. 2005. The Internet and the global monitoring of emerging diseases: lessons from the first 10 years of ProMED-mail. *Arch. Med. Res.* 36:724–30

51. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, et al. 2004. Implementing syndromic surveillance: a practical guide informed by the early experience. *J. Am. Med. Inform. Assoc.* 11:141–50

52. Massad E, Burattini MN, Lopez LF, Coutinho FAB. 2005. Forecasting versus projection models in epidemiology: the case of the SARS epidemics. *Med. Hypotheses* 65:17–22

53. Moore K, Black J, Rowe S, Franklin L. 2011. Syndromic surveillance for influenza in two hospital emergency departments. Relationships between ICD-10 codes and notified cases, before and during a pandemic. *BMC Public Health* 11:338

54. Moss R, Zarebski A, Dawson P, McCaw JM. 2017. Retrospective forecasting of the 2010–2014 Melbourne influenza seasons using multiple surveillance systems. *Epidemiol. Infect.* 145:156–69

55. Nouvellet P, Cori A, Garske T, Blake IM, Dorigatti I, et al. 2018. A simple approach to measure transmissibility and forecast incidence. *Epidemics* 22:29–35

56. Nsoesie EO, Brownstein JS, Ramakrishnan N, Marathe MV. 2014. A systematic review of studies on forecasting the dynamics of influenza outbreaks. *Influenza Other Respir. Viruses* 8:309–16

57. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. 2013. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLOS Comput. Biol.* 9:e1003256

58. Opatowski L, Fraser C, Griffin J, de Silva E, Van Kerkhove MD, et al. 2011. Transmission characteristics of the 2009 H1N1 influenza pandemic: comparison of 8 Southern hemisphere countries. *PLOS Pathog.* 7:e1002225

59. Osthus D, Hickmann KS, Caragea PC, Higdon D, Del Valle SY. 2017. Forecasting seasonal influenza with a state-space SIR model. *Ann. Appl. Stat.* 11:202–24

60. Patterson-Lomba O, Van Noort S, Cowling BJ, Wallinga J, Gomes MGM, et al. 2014. Utilizing syndromic surveillance data for estimating levels of influenza circulation. *Am. J. Epidemiol.* 179:1394–401

61. Perrotta D, Bella A, Rizzo C, Paolotti D. 2017. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in Italy. *PLOS ONE* 12:e0169801

62. Polgreen PM, Nelson FD, Neumann GR. 2007. Use of prediction markets to forecast infectious disease activity. *Clin. Infect. Dis.* 44:272–79

63. Preis T, Moat HS. 2014. Adaptive nowcasting of influenza outbreaks using *Google* searches. *R. Soc. Open Sci.* 1:140095

64. Priedhorsky R, Osthus D, Daughton AR, Moran KR, Generous N, et al. 2017. Measuring global disease with Wikipedia: success, failure, and a research agenda. *CSCW Conf. Comput. Support. Coop. Work* 2017:1812–34

65. Reich NG, Brooks LC, Fox SJ, Kandula S, McGowan CJ, et al. 2019. A collaborative multiyear, multi-model assessment of seasonal influenza forecasting in the United States. *PNAS* 116:3146–54

66. Ryu S, Ali ST, Cowling BJ, Lau EHY. 2020. Effects of school holidays on seasonal influenza in South Korea, 2014–2016. *J. Infect. Dis.* 222:832–35

67. Sanicas M, Forleo E, Pozzi G, Diop D. 2014. A review of the surveillance systems of influenza in selected countries in the tropical region. *Pan Afr. Med. J.* 19:121

68. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. 2015. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLOS Comput. Biol.* 11:e1004513

69. Santillana M, Zhang DW, Althouse BM, Ayers JW. 2014. What can digital disease detection learn from (an external revision to) Google Flu Trends? *Am. J. Prev. Med.* 47:341–47

70. Schmidt C. 2019. Real-time flu tracking. *Nature* 573:S58–59

71. Shaman J, Karspeck A. 2012. Forecasting seasonal outbreaks of influenza. *PNAS* 109:20425–30

72. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. 2013. Real-time influenza forecasts during the 2012–2013 season. *Nat. Commun.* 4:2837

73. Shaman J, Kohn M. 2009. Absolute humidity modulates influenza survival, transmission, and seasonality. *PNAS* 106:3243–48

74. Simonsen L, Gog JR, Olson D, Viboud C. 2016. Infectious disease surveillance in the big data era: towards faster and locally relevant systems. *J. Infect. Dis.* 214:S380–85

75. Simusika P, Tempia S, Chentulo E, Polansky L, Mazaba ML, et al. 2020. An evaluation of the Zambia influenza sentinel surveillance system, 2011–2017. *BMC Health Serv. Res.* 20:35

76. Smith PF, Hadler JL, Stanbury M, Rolfs RT, Hopkins RS, CSTE Surveill. Strategy Group. 2013. "Blueprint version 2.0": updating public health surveillance for the 21st century. *J. Public Health Manag. Pract.* 19:231–39

77. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, et al. 2015. Flu Near You: crowd-sourced symptom reporting spanning 2 influenza seasons. *Am. J. Public Health* 105:2124–30

78. Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, et al. 2013. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLOS Pathog.* 9:e1003194

79. Thompson LH, Malik MT, Gumel A, Strome T, Mahmud SM. 2014. Emergency department and 'Google Flu Trends' data as syndromic surveillance indicators for seasonal influenza. *Epidemiol. Infect.* 142:2397–405

80. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, et al. 2019. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics* 29:100356

81. Towers S, Feng Z. 2009. Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States. *Euro Surveill.* 14:19358

82. van de Kassteele J, Eilers PHC, Wallinga J. 2019. Nowcasting the number of new symptomatic cases during infectious disease outbreaks using constrained P-spline smoothing. *Epidemiology* 30:737–45

83. Viboud C, Boëlle PY, Carrat F, Valleron A-J, Flahault A. 2003. Prediction of the spread of influenza epidemics by the method of analogues. *Am. J. Epidemiol.* 158:996–1006

84. Viboud C, Vespignani A. 2019. The future of influenza forecasts. *PNAS* 116:2802–4

85. Waalen K, Kilander A, Dudman SG, Ramos-Ocao R, Hungnes O. 2012. Age-dependent prevalence of antibodies cross-reactive to the influenza A(H3N2) variant virus in sera collected in Norway in 2011. *Euro Surveill.* 17:20170

86. WHO (World Health Organ.). 2012. *World Health Organization interim global epidemiological surveillance standards for influenza (July 2012)*. Rep., WHO, Geneva. **https://www.who.int/influenza/resources/documents/INFSURVMANUAL.pdf**

87. WHO (World Health Organ.). 2020. *Influenza (seasonal). Ask the expert: influenza Q&A.* Influenza Fact Sheet, WHO, Geneva. **http://www.who.int/mediacentre/factsheets/fs211/en/index.html**

88. WHO (World Health Organ.). 2020. Into the history of influenza control. *Influenza.* **http://www.who.int/influenza/gip-anniversary/en/**

89. WHO (World Health Organ.). 2020. Public health surveillance. *Immunization, Vaccines and Biologicals.* **https://www.who.int/immunization/monitoring_surveillance/burden/vpd/en/**

90. Wong C-M, Chan K-P, Hedley AJ, Peiris JSM. 2004. Influenza-associated mortality in Hong Kong. *Clin. Infect. Dis.* 39:1611–17

91. Wong JY, Wu P, Nishiura H, Goldstein E, Lau EH, et al. 2013. Infection fatality risk of the pandemic A(H1N1)2009 virus in Hong Kong. *Am. J. Epidemiol.* 177:834–40

92. Wu P, Presanis AM, Bond HS, Lau EHY, Fang VJ, Cowling BJ. 2017. A joint analysis of influenza-associated hospitalizations and mortality in Hong Kong, 1998–2013. *Sci. Rep.* 7:929

93. Xu Q, Gel YR, Ramirez Ramirez LL, Nezafati K, Zhang Q, Tsui K-L. 2017. Forecasting influenza in Hong Kong with Google search queries and statistical model fusion. *PLOS ONE* 12:e0176690

94. Yamana TK, Kandula S, Shaman J. 2017. Individual versus superensemble forecasts of seasonal influenza outbreaks in the United States. *PLOS Comput. Biol.* 13:e1005801

95. Yang S, Santillana M, Brownstein JS, Gray J, Richardson S, Kou SC. 2017. Using electronic health records and Internet search information for accurate influenza forecasting. *BMC Infect. Dis.* 17:332
96. Yang W, Cowling BJ, Lau EH, Shaman J. 2015. Forecasting influenza epidemics in Hong Kong. *PLOS Comput. Biol.* 11:e1004383
97. Yang W, Olson DR, Shaman J. 2016. Forecasting influenza outbreaks in boroughs and neighborhoods of New York City. *PLOS Comput. Biol.* 12:e1005201
98. Zheng W, Aitken R, Muscatello DJ, Churches T. 2007. Potential for early warning of viral influenza activity in the community by monitoring clinical diagnoses of influenza in hospital emergency departments. *BMC Public Health* 7:250
99. Ziegler T, Mamahit A, Cox NJ. 2018. 65 years of influenza surveillance by a World Health Organization-coordinated global network. *Influenza Other Respir. Viruses* 12:558–65