

# A Smartphone-Driven Thermometer Application for Real-time Population- and Individual-Level Influenza Surveillance

Aaron C. Miller,<sup>1</sup> Inder Singh,<sup>2</sup> Erin Koehler,<sup>2</sup> and Philip M. Polgreen<sup>3</sup>

<sup>1</sup>Department of Computer Science, University of Iowa, Iowa City; <sup>2</sup>Kinsa, Inc, San Francisco, California; and <sup>3</sup>Departments of Internal Medicine and Epidemiology, Carver College of Medicine, University of Iowa, Iowa City

**Background.** Smartphone-based sensors may enable real-time surveillance of infectious diseases at population and household levels. This study evaluates the use of data from commercially available “smart thermometers,” connected to a mobile phone application, for surveillance of influenza-like illness (ILI).

**Methods.** At a population level, we analyzed the correlation between thermometer recordings and Centers for Disease Control and Prevention–reported ILI activity nationally and by age group and region. We developed time-series models to forecast ILI activity in real time and up to 3 weeks in advance. We analyzed the ability of thermometer readings to track the duration of fevers and identify biphasic fever patterns. We also investigated potential transmission of febrile illness within households among device users.

**Results.** Thermometer readings were highly correlated with national ILI activity ( $r > 0.95$ ) and activity patterns across regions and age groups. Thermometer readings also significantly improved forecasts of ILI activity in real time and up to 3 weeks in advance. We found that fevers lasting between 3 and 6 days and biphasic fever episodes occurred more frequently during the influenza season. In addition, potential cases of in-household transmission of febrile illness originated more frequently from children than adults.

**Conclusions.** Smart thermometers represent a novel source of information for influenza surveillance and forecasting. Thermometer readings capture real-time ILI activity at a population level, and they can also be used to generate improved forecasts. Moreover, the widespread deployment of these smart thermometers may also allow for more rapid and efficient surveillance at the household level.

**Keywords.** influenza; surveillance; seasonality; thermometer; smartphone.

Mobile devices are frequently used to inform patients and healthcare professionals. These devices can gather information from patients [1–3] and also be paired with sensors to monitor patients remotely. Examples of such sensors include scales and glucose or blood pressure monitors [4, 5]. Aggregated sensor data may also provide population-level information. Specifically, time-stamped geocoded health-sensor data may inform disease-surveillance efforts.

Influenza surveillance is especially important as influenza is a major cause of morbidity and mortality [6, 7]. Influenza also increases the risk for bacterial infections [8–11] and exacerbates chronic illnesses [12–15]. Although influenza outbreaks occur annually, their timing and amplitude vary considerably. Vaccinations and treatments exist, but their efficacy depends on timing. Even minimal advanced warning of increasing levels of

influenza activity can inform prevention-and-treatment efforts. Yet, influenza surveillance reports typically lag by 1–2 weeks.

Given the delay inherent with traditional surveillance approaches, various novel information sources have been proposed, including Internet search volume [16, 17], prediction markets [18], Twitter posts [19], Wikipedia views [20, 21], medication sales [22], and thermometer sales [23]. However, these data represent a proxy for clinical information. In contrast, geolocated fever measurements from “smart thermometers” (ie, thermometers paired to mobile devices) measure a clinical sign in real time directly. In addition, thermometer data joined with demographic and symptom data supplied by users may help characterize particular features of an influenza season or help supplement traditional epidemiological field work (eg, estimating household transmission risk) [24, 25].

The purpose of this article is to explore the utility of using large-scale geocoded data from commercially available smart thermometers to perform population- and individual-level influenza surveillance. At a population level, we demonstrate the ability to capture regional and age-based seasonal patterns and to generate national-level influenza forecasts. At an individual level, we demonstrate the promise of using thermometer-based data for epidemiology field work by monitoring fever

Received 25 September 2017; editorial decision 24 January 2018; accepted 29 January 2018; published online February 8, 2018.

Correspondence: P. M. Polgreen, Departments of Medicine and Epidemiology, University of Iowa, 200 Hawkins Dr, Iowa City, IA 52242 ([philip-polgreen@uiowa.edu](mailto:philip-polgreen@uiowa.edu)).

Clinical Infectious Diseases® 2018;67(3):388–97

© The Author(s) 2018. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: [journals.permissions@oup.com](mailto:journals.permissions@oup.com). DOI: 10.1093/cid/ciy073

duration, biphasic fever episodes, and potential within-household transmission of febrile episodes during the influenza season.

## METHODS

### Study Data

Kinsa Smart Thermometers record and store temperature measurements, using the Kinsa smartphone application, which is compatible with a range of mobile devices. The Kinsa Smart Thermometer was introduced in April 2013 to record oral, underarm, and rectal temperatures. The Kinsa Smart Ear Thermometer was released in November 2015. When recording temperatures, users can assign readings to profiles by age and sex, allowing readings from multiple users within a household to be distinguished. Readings are geocoded using Global Positioning System location (for enabled devices) or by Internet Protocol address. Because we used only de-identified data, the University of Iowa Institutional Review Board designated this study as nonhuman subjects research.

Our study period was 30 August 2015 to 23 December 2017. We observed a slight product uptake effect at the beginning of our study period and applied a correction to de-trend these data for this effect until the week of 7 February 2016, when the effect dissipated as described in Supplementary Appendix 1. Temperature readings were also filtered to only include values ranging from 34°C to 43°C (93.2°F –109.4°F).

To measure influenza activity, we used nationally and regionally weighted data for weekly influenza-like illness (ILI) collected by the Centers for Disease Control and Prevention (CDC) Outpatient ILI Surveillance Network (ILINet). ILINet measures the weekly percentage of outpatient physician visits with a “temperature 100°F [37.8°C] or greater and a cough and/or a sore throat without a known cause other than influenza” (<https://www.cdc.gov/flu/weekly/overview.htm>). The CDC makes occasional ILI data revisions, so we also collected data available at the time of reporting. [Supplementary Figure 1](#) provides a comparison of the originally released and corrected ILI data. We also make age-based comparisons for ILI activity by computing the percentage of all physician visits that were attributable to ILI-related visits in the age groups reported by the CDC. (Note: Because the CDC does not report the denominator for age-specific visit counts, aged-based ILI incidence has a slightly different interpretation.)

### Forecasting Population-Level Influenza Activity

We considered 3 primary ways to capture ILI activity using thermometer readings: (1) weekly counts of total temperature readings; (2) total fever readings (ie, temperature  $\geq 37.8^\circ\text{C}$  [ $\geq 100^\circ\text{F}$ ] to match the CDC ILINet definition); and (3) the total number of distinct fever episodes, identified by distinct user

profiles registering a fever in a given week. We compared each of these series to CDC-reported ILI activity at a national level and segmented regionally by age.

We analyzed the ability of thermometer readings to provide additional information in forecasting ILI activity, during current and future time periods. Because thermometer readings are available in real time, while ILI reports lag by approximately 2 weeks, we use temperature readings, at time  $t$ , to construct forecasts (referred to as “nowcasts”) of ILI activity at time  $t$ . In addition, we used readings, at time  $t$ , to forecast ILI activity at weeks  $t + 1$ ,  $t + 2$ , or  $t + 3$ . Models were constructed for both 1- and 2-week lags in ILI reporting. To generate estimates more representative of real-time forecasting, we used the preliminarily released ILI data from archived weekly CDC surveillance reports to train models near the point of prediction and to produce out-of-sample estimates. Model specifications are detailed in [Supplementary Appendix 2](#).

To evaluate forecasts, we used an adaptive out-of-sample forecasting approach [26–28]. We use a sliding 52-week training period to capture the dynamics of a complete influenza cycle, training each model on 52 weeks of data and evaluating predictions on the subsequent week. This procedure was repeated, iteratively, to generate 68 out-of-sample forecasts. To analyze prediction improvement offered by each thermometer based-variable, we compared a baseline autoregressive integrated moving average (ARIMA) model to an extended ARIMA (ARIMAX) model incorporating the exogenous thermometer data. The ARIMA model uses lagged ILI values, while the ARIMAX model incorporates lagged ILI data along with current and lagged values of the thermometer data. To provide an out-of-sample comparison, we used an automated selection process to select the model order for our ARIMA/ARIMAX models, based on the model fit in the 52-week training period. To analyze forecasting performance, we used a generalized linear model with autoregressive components for ILI. [Supplementary Appendix 3](#) provides additional details on model selection.

### Individual-Level Influenza Surveillance

To study individual-level fever patterns, we only considered readings in those cases in which users provided complete profile information. We identified distinct fever episodes associated with each user profile, defined as a period of 1 or more consecutive days during which a fever was sustained. We identified distinct fever episodes by finding the point at which a fever first occurred and linking across days for which consecutive fevers were recorded. We grouped fever recordings into a single episode if they were separated by 1 or 2 days with no readings in between. Fevers separated by  $\geq 3$  days without a temperature reading, or a day with only nonfever temperature readings in between, were treated as separate episodes.

### Measuring Seasonal Trends in Fever Duration

First, we computed the correlation between average fever duration and ILI activity. Second, because many febrile episodes had only 1 reading, we performed an additional set of analyses focused on the number of distinct days for each fever episode. We performed 2 analyses to relate fever duration to influenza activity. We created subsets of episodes in cohorts based on the days of duration (eg, 2 days, 3 days) and computed the correlation between the relative frequency of a given fever-duration cohort each week and the ILI series. Also, for each fever-duration cohort, we computed the odds ratio that fever episodes of a specific duration would occur during the 6-week peak of influenza activity relative to the 6-week nadir.

### Measuring Seasonal Trends of Biphasic Fever Episodes

Influenza can be associated with a syndrome characterized by a biphasic fever when there is a subsequent secondary infection (eg, pneumonia) after a recovery from influenza [8, 9, 11, 29]. We identified biphasic fever episodes as consecutive fever episodes in a single user that meet 1 of 2 criteria: The episodes were separated by 4–7 days with no readings in between, or were separated by <7 days and had at least 1 day containing only fever-free temperature readings in between. Next, we compared the total number and percentage of weekly biphasic fever episodes to ILI activity.

### Detecting Possible Transmission Events

To analyze the frequency of potential in-household disease transmission of febrile illnesses, we identified distinct fever episodes recorded by devices with multiple user profiles. Two fever episodes, recorded by a single device, were labeled as a potential transmission event if they meet the following criteria: start dates of the 2 episodes occurred on different days, and the start date of the second episode is within periods of 3, 5, or 7 days of the preceding episode. We then calculated the number of potential transmissions that occurred from child-to-adult, child-to-child, adult-to-adult, or adult-to-child. Finally, we computed the relative likelihood of a transmission event originating with a child vs an adult.

## RESULTS

### Device Coverage and Temperature Reporting

From 30 August 2015 to 23 December 2017, there were 8 234 027 temperature readings generated by 448 321 devices. Of these devices, 362 030 had at least 1 user profile identified by user age and sex, for a total of 604 982 different profiles. Because users may not enter a profile, approximately half of the observations (3461 117) did not have a value for age or sex.

Temperature readings were reported in all 50 states. [Supplementary Figure 2](#) displays per-capita readings and devices by state. [Table 1](#) provides a count of user profiles by age and sex. The majority of user profiles reflect an age distribution consistent with children and parents. There were a similar number of male and female children and nearly twice as many female as male adults.

### Correlation Between Thermometer Readings and ILI Activity

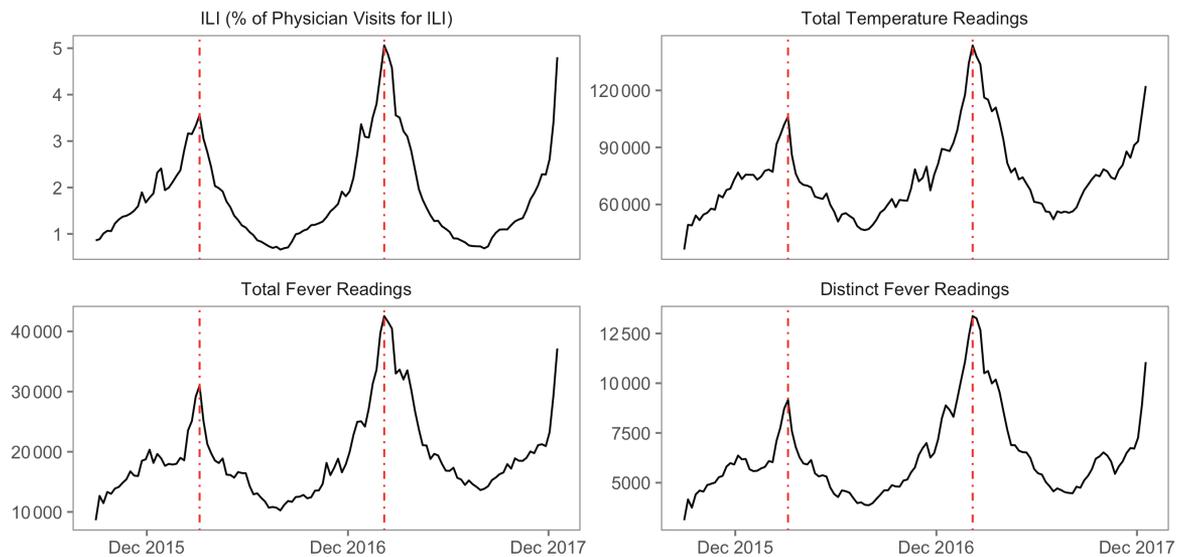
Thermometer counts were highly correlated with national ILI activity. [Figure 1](#) depicts the national-level trends in ILI, along with the total number of weekly readings, fever readings, and distinct fever events. The correlation coefficient between ILI and total readings, total fever readings, and distinct fevers were .938, .940, and .928, respectively. Because distinct fevers represented the closest approximation for the number of users experiencing a fever each week, we report the remaining results using this value. Results using total readings and total counts are reported in the [Supplementary Materials](#).

Thermometer data segmented by geographic location and age were compared to ILI activity in corresponding locations and age groups. [Figures 2](#) and [3](#) compare ILI activity and the count of distinct fevers by age group and by CDC region, respectively. Across age groups, the correlation between distinct fevers and ILI activity ranged from .784 for age ≥65 years, to .979 for age 25–49 years. Across CDC regions, the correlation ranged from .704 in region 9 (Arizona, California, Guam, Hawaii, Nevada), to .942 in region 5 (Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin). [Supplementary Tables 1](#) and [2](#) contain a summary of correlation coefficients, and [Supplementary Appendix 4](#) provides a more detailed analysis.

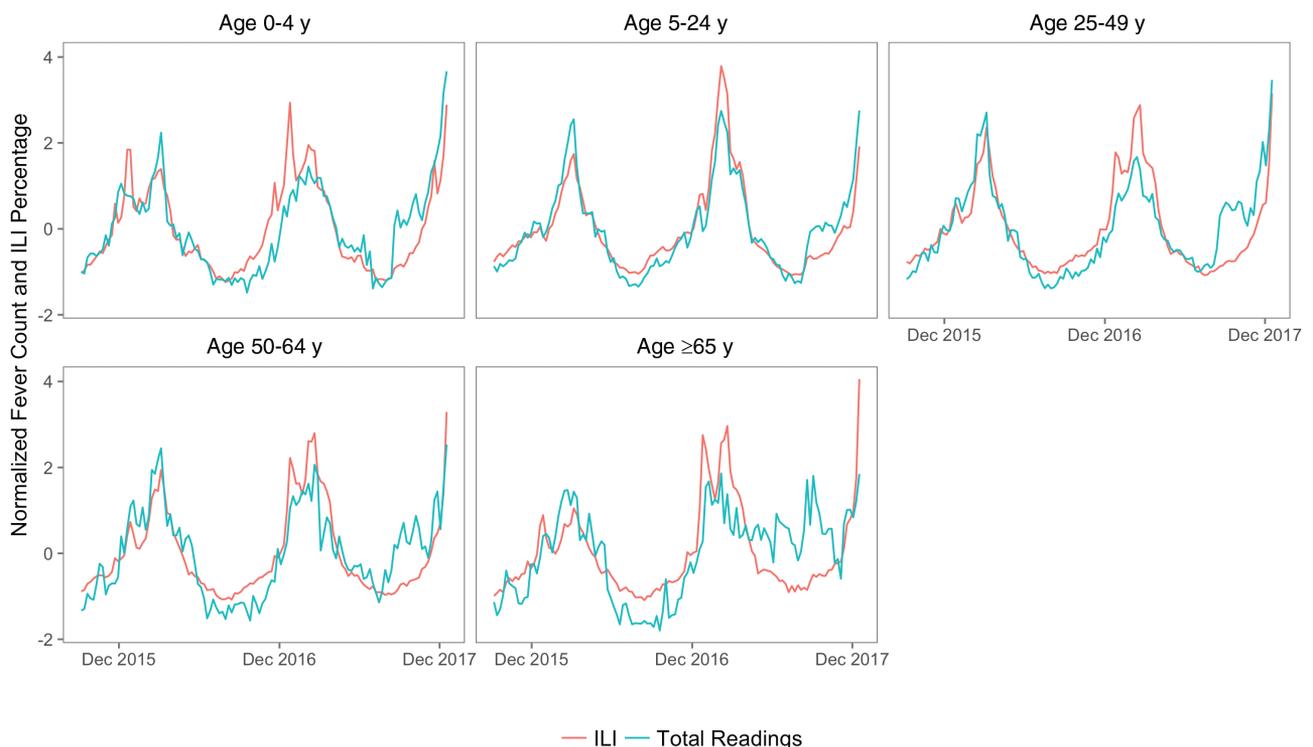
**Table 1. Counts of Smart Thermometer Users and Readings by Age and Sex**

Age Group, y	Female		Male		Unknown Sex	
	Readings	Users	Readings	Users	Readings	Users
0–4	632 747	69 813	663 609	72 036	204 893	34 395
5–24	923 371	113 058	733 828	95 016	71 819	10 551
25–49	823 337	103 820	295 587	59 227	46 959	7384
50–64	143 270	15 354	104 149	12 027	11 051	1339
≥65	43 436	4469	53 733	4452	4256	437
Unknown	6567	452	10 298	442	3461 117	

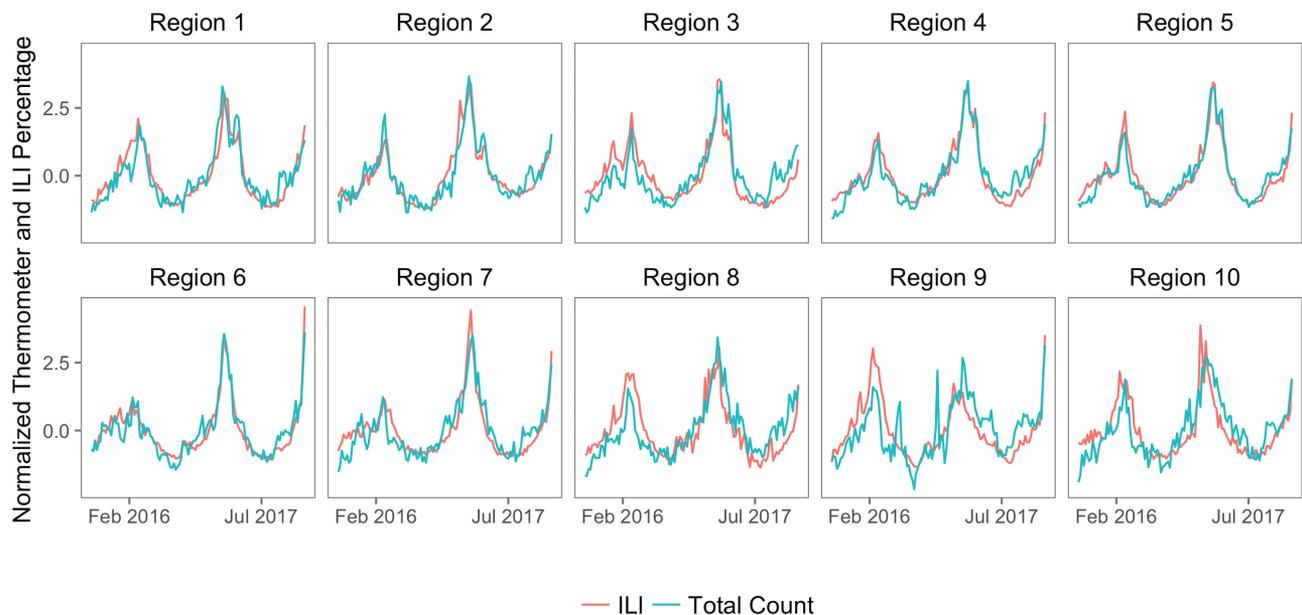
For readings where no sex or age was provided, we were unable to derive the total number of users.



**Figure 1.** Physician visits for influenza-like illness (ILI) along with counts of total temperature readings, total fever readings, and distinct fever episodes are plotted by week. The dashed red line marks the peak of influenza seasons, as determined by ILI visits. All 3 thermometer-based counts (total weekly readings, total fever readings, distinct fevers) follow the pattern in ILI activity, with each peak co-occurring with peaks in ILI activity. All thermometer-based series were corrected to account for product uptake. This correction was applied from the start of the period until 17 February 2016, as described in Supplementary Appendix 1.



**Figure 2.** Weekly influenza-like illness (ILI) activity and counts of distinct fever episodes are broken down by age group, normalized and plotted. All series were normalized by their respective mean and standard deviation. Because the Centers for Disease Control and Prevention (CDC) does not report age-specific ILI incidence, we have computed ILI incidence by dividing the CDC-reported counts of ILI-related visits in each age group by the total number of reported outpatient visits. The CDC does not provide age-specific total visit counts, so these values should be interpreted differently than other ILI incidence. Fever readings for each age group follows the trend in ILI activity for the corresponding age group. Pearson correlation coefficients between ILI activity and distinct fever events were .881, .938, .979, .906, and .784, for ages 0–4, 5–24, 25–49, 50–64, and ≥65 years, respectively. Age groups 5–24 years and 25–49 years are most well represented in the thermometer data, whereas age ≥50 years had far fewer device readings (see Table 1). The lack of device coverage likely explains the additional noise in these age groups.



**Figure 3.** Weekly influenza-like illness (ILI) activity and counts of distinct fever events are broken into regions, normalized and plotted. All series were normalized by their respective mean and standard deviation. Pearson correlation coefficients between ILI activity and distinct fever counts were .875, .928, .847, .924, .942, .909, .888, .759, .704, and .725, for regions 1–10, respectively. Centers for Disease Control and Prevention regions are defined at: <https://www.cdc.gov/coordinatedchronic/docs/NCCDPHP-Regions-Map.pdf>. Our region 1 corresponds to their region A, etc.

### Forecasting Population-Level Influenza Activity

Table 2 reports out-of-sample nowcast and forecast performance for baseline models and models incorporating distinct fever counts (Supplementary Table 3 reports results for all thermometer-based counts). Using a 2-week lag in ILI activity, fever counts provided a reduction in out-of-sample forecast error of around 35.1%–53.2% and increased out-of-sample forecast correlation by >10%. Figure 4 plots the out-of-sample nowcasts for the different thermometer variables, using either a 1- or 2-week lag in ILI. We also analyzed the performance of thermometer

data in forecasting 1, 2, and 3 weeks ahead of the current time period. When using a 2-week lag in ILI reporting, for example, the count of distinct fevers reduced out-of-sample error by 28.8%–38.0%, 24.5%–27.9%, and 8.3%–14.4% in the 1-, 2-, and 3-week forecasts, respectively.

### Individual-Level Influenza Surveillance

#### Fever Duration

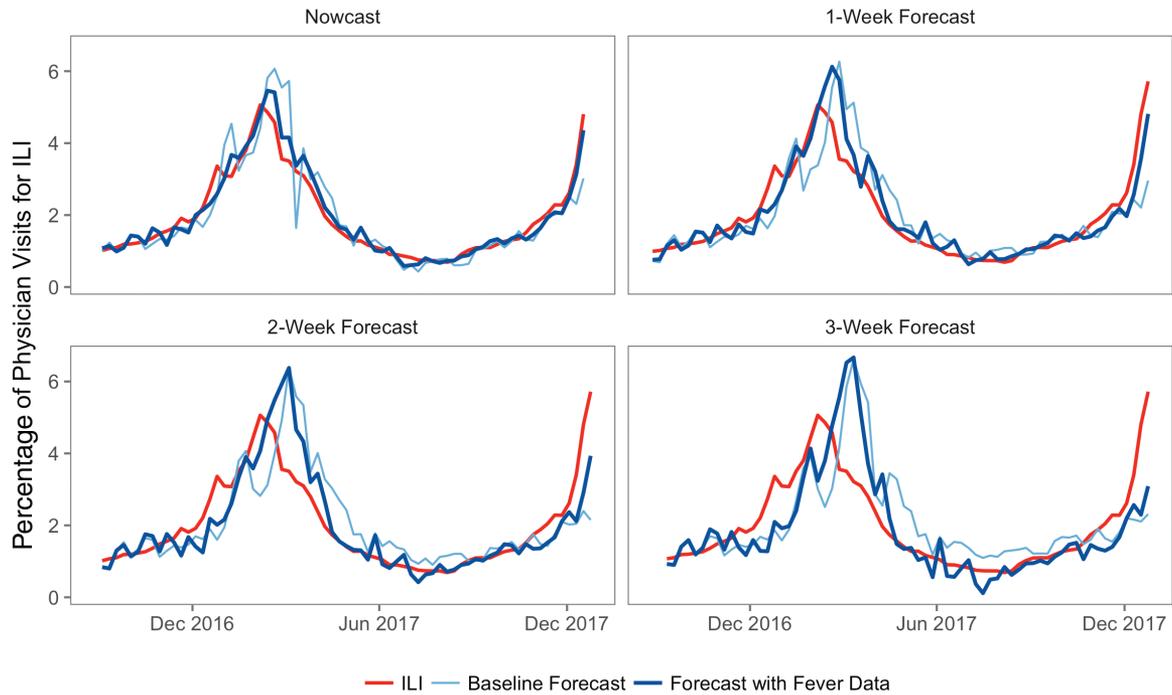
Of 710620 distinct fever episodes, 484878 had user profile information. The average duration of fevers by week

**Table 2. Out-of-Sample Nowcast and Forecast Results Using Counts of Distinct Fevers (Percentage Improvement Compared to Baseline)**

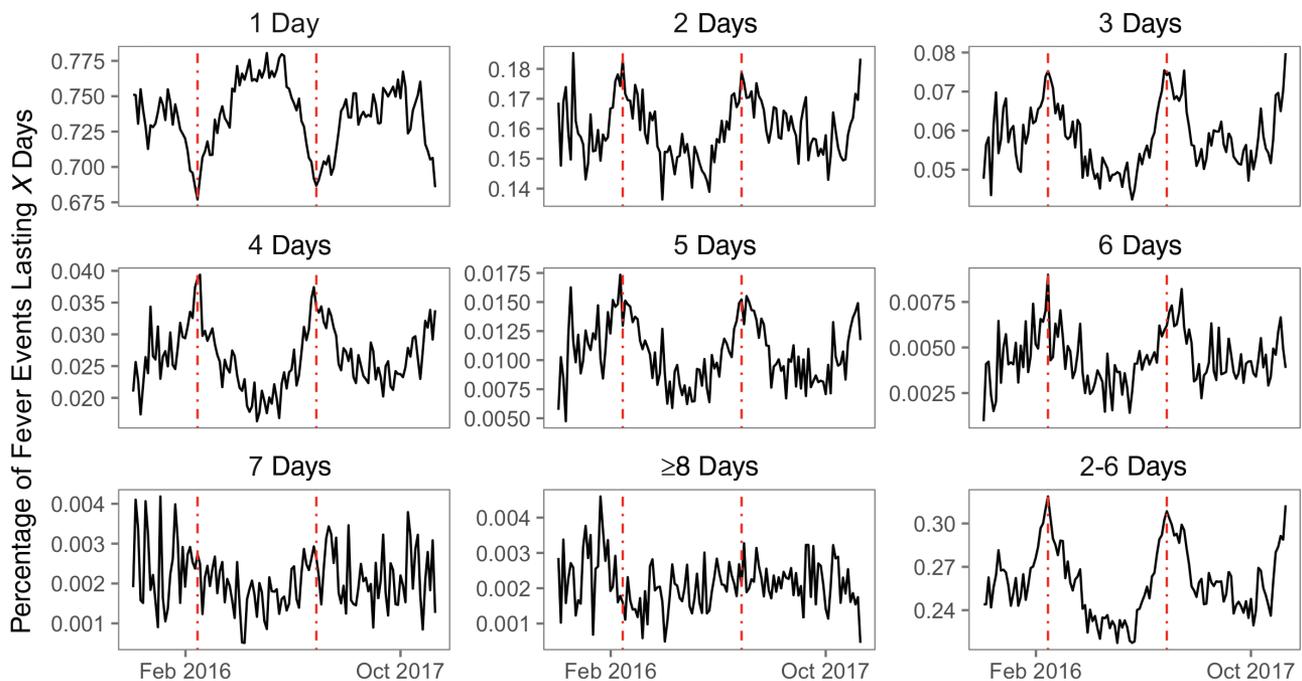
	2-Week ILI Reporting Lag				1-Week ILI Reporting Lag			
	RMSE	MAPE	MAE	<i>r</i>	RMSE	MAPE	MAE	<i>r</i>
<b>Nowcast</b>								
Baseline	0.648	18.9	0.419	0.880	0.357	11.0	0.224	0.960
Distinct fevers	0.303 (53.2)	12.3 (35.1)	0.229 (45.3)	0.968 (10.0)	0.220 (38.5)	8.2 (25.1)	0.151 (32.6)	0.985 (2.6)
<b>1-week forecast</b>								
Baseline	0.825	25.4	0.555	0.819	0.569	17.9	0.392	0.910
Distinct fevers	0.512 (38.0)	18.1 (28.8)	0.364 (34.5)	0.944 (15.3)	0.410 (28.0)	13.7 (23.2)	0.281 (28.3)	0.949 (4.3)
<b>2-week forecast</b>								
Baseline	1.018	32.6	0.692	0.700	0.836	25.4	0.561	0.815
Distinct fevers	0.735 (27.9)	24.6 (24.5)	0.509 (26.4)	0.876 (25.0)	0.637 (23.8)	18.7 (26.4)	0.412 (26.5)	0.879 (7.8)
<b>3-week forecast</b>								
Baseline	1.147	39.5	0.802	0.596	1.028	32.8	0.7	0.694
Distinct fevers	1.052 (8.3)	33.9 (14.4)	0.717 (10.7)	0.741 (24.4)	0.862 (16.2)	27.7 (15.6)	0.585 (16.4)	0.802 (15.5)

The percentage improvement is shown in parentheses for each extended autoregressive integrated moving average model, containing thermometer data, compared to the baseline autoregressive integrated moving average model, without thermometer data. Performance metrics are provided for RMSE, MAPE, MAE, and Pearson correlation (*r*).

Abbreviations: ILI, influenza-like illness; MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean squared error.



**Figure 4.** Out-of-sample nowcasts and forecasts for influenza-like illness (ILI) activity using distinct fever events are plotted against actual Centers for Disease Control and Prevention–reported ILI activity. For each forecasting window, baseline forecasts, using only lagged values of ILI, are plotted in light blue along with forecasts incorporating thermometer-based fever counts. Forecasts are highly correlated with ILI activity. Forecasts incorporating distinct fever episodes produced better out-of-sample estimates relative to the baseline model. Specifically, forecasts appeared more stable compared to the baseline model.



**Figure 5.** The percentage of fever events each week of a given duration are plotted for various duration periods. The dashed red line denotes influenza-like illness (ILI) peak in Centers for Disease Control and Prevention–reported ILI activity. The percentage of fevers each week that last between 2 and 6 days appear to match the trend in ILI activity. The percentage of fevers each week that last only 1 day is inversely related to the ILI trend, whereas fevers lasting  $\geq 8$  days do not exhibit a seasonal trend. The inverse relationship for fevers of 1 day's duration implies that fevers lasting  $>1$  day were relatively more common during flu season than outside of flu season.

**Table 3. Correlation Between Trends in Influenza-like Illness Activity and the Percentage of Fevers Each Week Lasting a Given Duration**

Correlation	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days	≥8 Days	3–6 Days
Series correlation with ILI <sup>a</sup>	-.742	.522	.700	.753	.704	.623	.272	.056	.780
Odds ratio (peak vs nadir) <sup>b</sup>									
2015–2016	0.68 (.66–.71)	1.21 (1.15–1.26)	1.45 (1.35–1.56)	1.75 (1.57–1.95)	1.94 (1.64–2.30)	2.40 (1.84–3.15)	1.92 (1.29–2.86)	1.03 (.71–1.49)	1.45 (1.40–1.51)
2016–2017	0.80 (.77–.82)	1.10 (1.06–1.15)	1.25 (1.18–1.33)	1.41 (1.30–1.54)	1.53 (1.34–1.76)	1.56 (1.27–1.92)	1.36 (.99–1.88)	1.21 (.88–1.65)	1.25 (1.21–1.29)

Abbreviation: ILI, influenza-like illness.

<sup>a</sup>Series correlation with ILI corresponds to the correlation between the weekly percentage of fevers of a given duration (ie, the series depicted in Figure 5) and weekly ILI activity.

<sup>b</sup>Odds ratios are calculated between the peak and nadir of flu season based on the number of fevers each week lasting a given duration, compared to any other duration. Confidence intervals, corresponding to a 95% level, are reported in parentheses. Peak and nadir were defined as the 6-week period around the ILI peak and nadir at the following dates: 2015–2016 peak: 7 February 2016 to 6 March 2016; 2015–2016 nadir: 10 July 2016 to 14 August 2016; 2016–2017 peak: 22 January 2017 to 26 February 2017; 2016–2017 nadir: 18 June 2017 to 23 July 2017.

was strongly correlated ( $r = 0.768$ ) with ILI activity, indicating that longer-lasting fevers are correlated with ILI activity (Supplementary Figure 3). Figure 5 depicts the percentage of fevers each week by days of duration. The weekly frequency of fevers lasting 2–6 days followed a pattern similar to ILI activity. Fevers lasting only 1 day were inversely related to ILI activity. Fevers >6 days exhibited no seasonality. Table 3 reports correlation coefficients between ILI and the percentage of weekly fevers of different duration. The percentage of fevers lasting between 3 and 6 days appeared to be most strongly correlated with ILI activity. Table 3 also reports odds ratios for the likelihood of fevers of a given duration occurring during the peak vs the nadir of influenza season. The odds of a fever lasting between 3 and 6 days were between 1.25 and 2.40 times more likely to occur during the peak of influenza season compared to the nadir.

#### Biphasic Fever Episodes

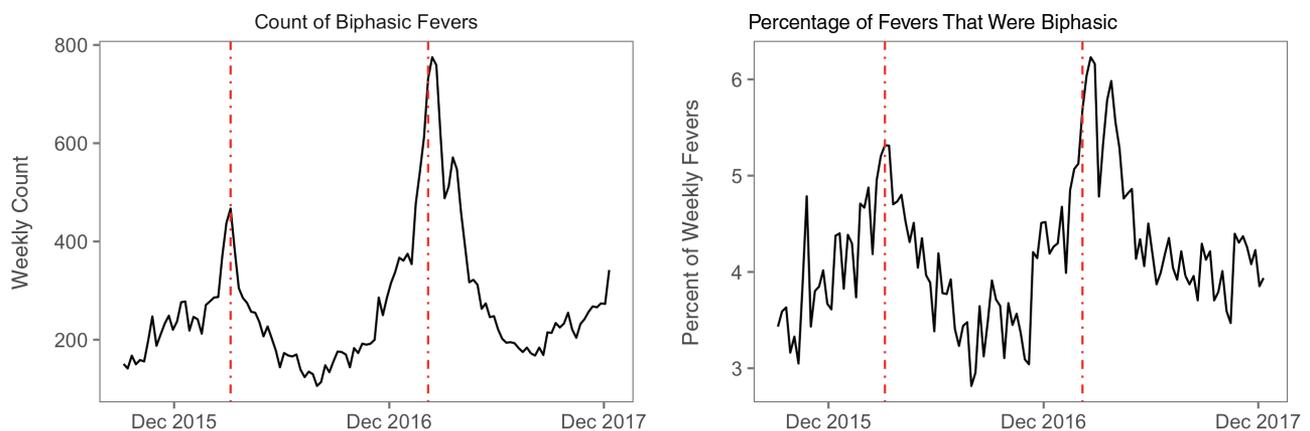
We identified a total of 31 270 biphasic fever episodes. Figure 6 shows that weekly counts and weekly percentage of biphasic fevers both reflected ILI activity. Counts and weekly percentages of biphasic fevers had a correlation with ILI of .857 and .773, respectively. The frequency of biphasic fevers ranged from around 3% to >6% from the nadir to the peak of influenza season.

#### Possible Transmission Episodes

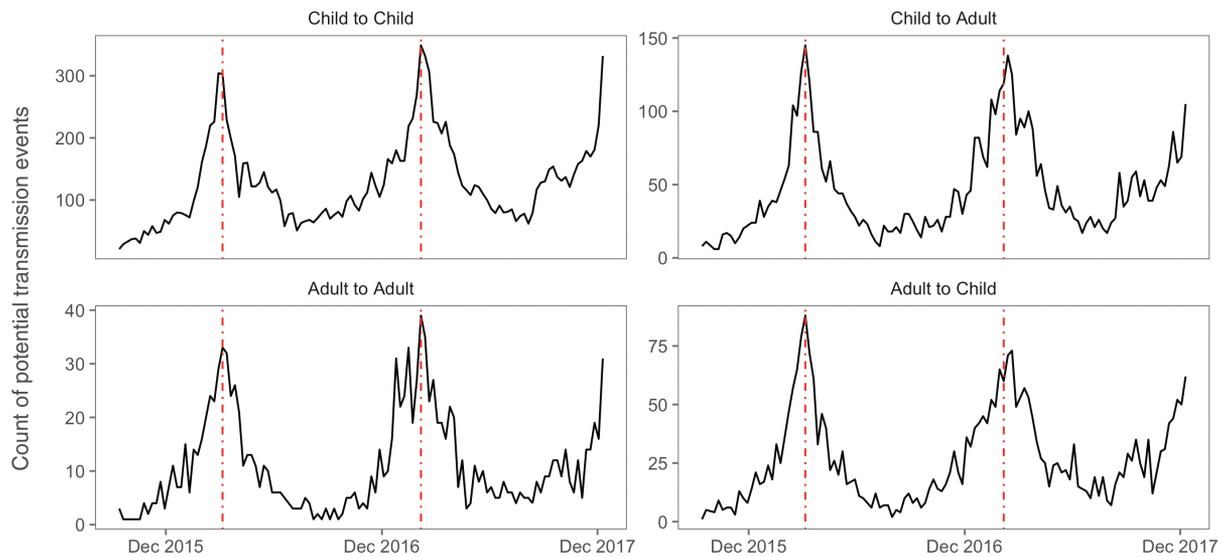
We identified 156 529 devices that had >1 distinct user profile, of which 63 320 devices had >1 profile reporting a fever episode during the study period. Of these devices, we identified a total of 50 305 fever episodes occurring 1–7 days after a fever episode in another device user, representing potential disease transmission. Figure 7 depicts potential weekly transmission events between children and adults, which were highly correlated with ILI. Table 4 describes the counts of potential fever transmissions. Across the various transmission windows, child-to-child transmission was most common, followed by child-to-adult, adult-to-child, and adult-to-adult, respectively. Depending on the transmission window, there were 4.41–5.08 times as many potential transmissions originating from a child than from an adult.

#### DISCUSSION

Temperature data aggregated from commercially available smart thermometers can capture influenza activity in real time nationally, regionally, and for different age groups. Forecasts from time-series models showed significant improvement when thermometer data were incorporated. Using de-identified user-profile data, we were also able to capture other clinical



**Figure 6.** Centers for Disease Control and Prevention–reported influenza-like illness (ILI) activity, the total number of biphasic fever events, and the percentage of fevers that are biphasic are plotted by week. Both the weekly totals and weekly percentage of biphasic fever events appear to follow the trend in ILI activity.



**Figure 7.** Counts of potential transmission events broken down by child or adult transmission. Transmission events are defined by a period of 1–7 days between fever start dates. All 4 types of transmission are highly correlated with influenza-like illness activity: child to child (.855), child to adult (.882), adult to adult (.877), and adult to child (.861).

features of febrile episodes: the duration of fevers, the incidence of biphasic fevers, and the frequency of potential transmission events. The correlation between our individual-level surveillance results and the known characteristics of influenza in conjunction with our population-level device provides supportive evidence that individual-level data may be useful for refining influenza surveillance approaches. Moreover, our ability to track individual-level characteristics of febrile episodes via user-generated data demonstrates the potential for performing field epidemiology work using mobile devices.

With only 2.5 years of data, simple linear models produced encouraging results for forecasting ILI in advance of CDC reports. Even assuming only a 1-week lag for the release of ILI data, we saw substantial improvements in out-of-sample nowcasts. In addition, our forecasting results may improve by incorporating more data and more sophisticated modeling approaches. Many of the additional clinical characteristics of

febrile episodes, captured by thermometer data, also tracked ILI activity (eg, patterns of fever duration, biphasic fever episodes, fevers among specific age groups, fevers in geographic regions) and should be evaluated for their potential to improve influenza activity forecasts in future work.

Using thermometer-based data to estimate influenza activity provides many potential benefits. Prior surveillance approaches have used data from clinical visits [30–33], but symptoms occur prior to visits, and thermometers can capture information before visits. Other efforts have analyzed purchases of over-the-counter medications [34, 35], telephone-triage calls [36], school-absentee data [37], or data from Internet searches or social media [16, 17, 19–21]. These data emerge prior to health-care visits, but represent proxy measures whereas thermometers capture an actual clinical sign. Finally, de-identified user-profile data provide additional demographic and clinical information about febrile episodes.

**Table 4. Counts of Potential In-household Transmission of Fever (Percentage of All Potential Transmissions Identified)**

Transmission	Transmission Window (Days Between the Start of Fever Episodes in Different Users) <sup>a</sup>		
	1–3 Days	1–5 Days	1–7 Days
Child to child	10 579 (56.4)	14 055 (57.3)	16 208 (57.2)
Child to adult	4014 (21.4)	5247 (21.4)	6093 (21.5)
Adult to child	2269 (12.1)	2684 (10.9)	3033 (10.7)
Adult to adult	1042 (5.6)	1233 (5.0)	1360 (4.8)
Indeterminate <sup>b</sup>	839 (4.5)	1322 (5.4)	1643 (5.8)
Child relative to adult transmission <sup>c</sup>	4.41	4.93	5.08

<sup>a</sup>The counts represented here use a transmission window defined as the start date between 2 consecutive fever episodes in 2 device users. However, because fever episodes frequently last >1 day, transmission windows can be defined based on the time between points within fever episodes. [Supplementary Table 4](#) provides expanded counts of fever episodes using a more sensitive definition to identify potential transmission events.

<sup>b</sup>Transmissions are indeterminate if both a prior child and adult fever occurred within the potential transmission window.

<sup>c</sup>Because children are overrepresented in the study data (see [Table 1](#)), these values simply reflect relative counts of potential transmission events and do not directly reflect risk or attack rates.

The ability of the thermometer's mobile app to track users via different profiles facilitates collection of information that is difficult, or impossible, to collect using traditional surveillance approaches. For example, influenza is commonly associated with a fever lasting multiple days, typically 3 days and up to 1 week [8, 29], and we were able to demonstrate that fevers of such duration were highly correlated with influenza seasons. We also found that biphasic fever episodes, another clinical phenomenon associated with influenza [8, 29], were highly correlated with influenza activity. Finally, we could track the direction of the spread of febrile illness from children to adults during influenza season, a pattern described in prior work [38, 39]. A similar approach could be used to estimate household transmission rates for influenza or other febrile illnesses.

Thermometer-based data provide many opportunities for future work. For example, broadcasting time- and location-specific alerts provides an opportunity to build personalized public health interventions (eg, vaccination reminders). The thermometer's app can currently input clinical symptoms (eg, cough, diarrhea), diagnoses, and medication reminders. This information could be used to refine surveillance approaches and perhaps help differentiate influenza from other febrile illnesses. Finally, if the app were to ask users about influenza vaccination status, we could possibly provide early estimates of vaccine effectiveness, especially if influenza cases could be confirmed. A recent study in China used a user-driven mobile health application to collect information on fevers and immunization status [40].

Despite promising results, our work has limitations. First, with only 2 years of data, our inferential conclusions are limited, and real-time performance may differ in future applications. Third, temperature readings may not uniformly cover socioeconomic or age groups or geographic locations. Increased product adoption, or efforts to increase device usage in underrepresented populations (eg, age >50 years) and regions, may lead to even more promising results. However, future work relies on continued product use. Finally, fevers are caused by many different infections. Future work should explore ways to confirm the cause of febrile episodes.

Smart thermometer-based data represent a timely and accurate source for surveillance of influenza. As data collected from these devices grow, and as more sophisticated modeling approaches are applied, we expect to provide even more accurate and longer-horizon forecasts. Moreover, given that the mobile application can collect information beyond temperature readings (eg, symptoms), there is an unparalleled opportunity to perform participatory research and field epidemiology for both established and emerging infectious diseases.

### Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

### Notes

**Acknowledgment.** I.S. and E.K. conceived and designed the Kinsa products for purposes of tracking the spread of illness, and supplied the de-identified data used in the analysis presented in this article. All other authors report no potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**Potential conflicts of interest.** I. S. and E. K. are employees of Kinsa, Inc, and are also shareholders in this company. They conceived and designed the Kinsa products for purposes of tracking the spread of illness, and supplied the de-identified data used in the analysis presented in this article. All other authors report no potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References

1. Chan YY, Wang P, Rogers L, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. *Nat Biotechnol* **2017**; 35:354–62.
2. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* **2012**; 12:67.
3. Boulton MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. *Biomed Eng Online* **2011**; 10:24.
4. Thomas JG, Bond DS. Review of innovations in digital health technology to promote weight control. *Curr Diab Rep* **2014**; 14:485.
5. Logan AG, McIsaac WJ, Tisler A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. *Am J Hypertens* **2007**; 20:942–8.
6. Thompson WW, Shay DK, Weintraub E, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA* **2003**; 289:179–86.
7. Thompson WW, Shay DK, Weintraub E, et al. Influenza-associated hospitalizations in the United States. *JAMA* **2004**; 292:1333–40.
8. Paules C, Subbarao K. Influenza. *Lancet* **2017**; 390:697–708.
9. Dawood FS, Chaves SS, Perez A, et al. Complications and associated bacterial coinfections among children hospitalized with seasonal or pandemic influenza, United States, 2003–2010. *J Infect Dis* **2014**; 209:686–94.
10. Heikkinen T, Silvennoinen H, Peltola V, et al. Burden of influenza in children in the community. *J Infect Dis* **2004**; 190:1369–73.
11. Chertow DS, Memoli MJ. Bacterial coinfection in influenza: a grand rounds review. *JAMA* **2013**; 309:275–82.
12. Foster ED, Cavanaugh JE, Haynes WG, et al. Acute myocardial infarctions, strokes and influenza: seasonal and pandemic effects. *Epidemiol Infect* **2013**; 141:735–44.
13. Gerke AK, Tang F, Yang M, Foster ED, Cavanaugh JE, Polgreen PM. Predicting chronic obstructive pulmonary disease hospitalizations based on concurrent influenza activity. *COPD* **2013**; 10:573–80.
14. Gerke AK, Yang M, Tang F, Foster ED, Cavanaugh JE, Polgreen PM. Association of hospitalizations for asthma with seasonal and pandemic influenza. *Respirology* **2014**; 19:116–21.
15. Warren-Gash C, Bhaskaran K, Hayward A, et al. Circulating influenza virus, climatic factors, and acute myocardial infarction: a time series study in England and Wales and Hong Kong. *J Infect Dis* **2011**; 203:1710–8.
16. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* **2009**; 457:1012–4.
17. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis* **2008**; 47:1443–8.
18. Polgreen PM, Nelson FD, Neumann GR. Use of prediction markets to forecast infectious disease activity. *Clin Infect Dis* **2007**; 44:272–9.
19. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* **2011**; 6:e19467.
20. Hickmann KS, Fairchild G, Priedhorsky R, et al. Forecasting the 2013–2014 influenza season using Wikipedia. *PLoS Comput Biol* **2015**; 11:e1004239.
21. McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS Comput Biol* **2014**; 10:e1003581.
22. Liu TY, Sanders JL, Tsui FC, Espino JU, Dato VM, Suyama J. Association of over-the-counter pharmaceutical sales with influenza-like-illnesses to patient volume in an urgent care setting. *PLoS One* **2013**; 8:e59273.
23. Villamarín R, Cooper G, Wagner M, Tsui FC, Espino JU. A method for estimating from thermometer sales the incidence of diseases that are symptomatically similar to influenza. *J Biomed Inform* **2013**; 46:444–57.

24. Tsang TK, Lau LL, Cauchemez S, Cowling BJ. Household transmission of influenza virus. *Trends Microbiol* **2016**; 24:123–33.
25. Lau LL, Nishiura H, Kelly H, Ip DK, Leung GM, Cowling BJ. Household transmission of 2009 pandemic influenza A (H1N1): a systematic review and meta-analysis. *Epidemiology* **2012**; 23:531–42.
26. Preis T, Moat HS. Adaptive nowcasting of influenza outbreaks using Google searches. *R Soc Open Sci* **2014**; 1:140095.
27. Santillana M, Zhang DW, Althouse BM, Ayers JW. What can digital disease detection learn from (an external revision to) Google Flu Trends? *Am J Prev Med* **2014**; 47:341–7.
28. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc Natl Acad Sci U S A* **2015**; 112:14473–8.
29. Bennett JE, Dolin R, Blaser MJ. Principles and practice of infectious diseases. Elsevier Health Sciences, **2014**.
30. Irvin CB, Nouhan PP, Rice K. Syndromic analysis of computerized emergency department patients' chief complaints: an opportunity for bioterrorism and influenza surveillance. *Ann Emerg Med* **2003**; 41:447–52.
31. Yuan CM, Love S, Wilson M. Syndromic surveillance at hospital emergency departments—southeastern Virginia. *MMWR Suppl* **2004**; 53:56–8.
32. Suyama J, Sztajnkrzyer M, Lindsell C, Otten EJ, Daniels JM, Kressel AB. Surveillance of infectious disease occurrences in the community: an analysis of symptom presentation in the emergency department. *Acad Emerg Med* **2003**; 10:753–63.
33. Olson DR, Heffernan RT, Paladini M, Konty K, Weiss D, Mostashari F. Monitoring the impact of influenza by age: emergency department fever and respiratory complaint surveillance in New York City. *PLoS Med* **2007**; 4:e247.
34. Welliver RC, Cherry JD, Boyer KM, et al. Sales of nonprescription cold remedies: a unique method of influenza surveillance. *Pediatr Res* **1979**; 13:1015–7.
35. Davies GR, Finch RG. Sales of over-the-counter remedies as an early warning system for winter bed crises. *Clin Microbiol Infect* **2003**; 9:858–63.
36. Espino JU, Hogan WR, Wagner MM. Telephone triage: a timely data source for surveillance of influenza-like diseases. *AMIA Annu Symp Proc* **2003**: 215–9.
37. Lenaway DD, Ambler A. Evaluation of a school-based influenza surveillance system. *Public Health Rep* **1995**; 110:333–7.
38. Longini IM Jr, Koopman JS, Monto AS, Fox JP. Estimating household and community transmission parameters for influenza. *Am J Epidemiol* **1982**; 115:736–51.
39. Glezen WP, Couch RB. Interpandemic influenza in the Houston area, 1974–76. *N Engl J Med* **1978**; 298:587–92.
40. Hsuen Y, Brownstein JS, Liu J, Hawkins JB. Use of a digital health application for influenza surveillance in China. *Am J Public Health* **2017**; 107:1130–6.