Accuracy of State-Level Surveillance during Emerging Outbreaks of Respiratory Viruses: A Model-Based Assessment

Yuwen Gu, Elise DeDoncker, Richard VanEnk, Rajib Paul, Susan Peters, Gillian Stoltman, and Diana Prieto

It is long perceived that the more data collection, the more knowledge emerges about the real disease progression. During emergencies like the H1N1 and the severe acute respiratory syndrome coronavirus 2 pandemics, public health surveillance requested increased testing to address the exacerbated demand. However, it is currently unknown how accurately surveillance portrays disease progression through incidence and confirmed case trends. State surveillance, unlike commercial testing, can process specimens based on the upcoming demand (e.g., with testing restrictions). Hence, proper assessment of accuracy may lead to improvements for a robust infrastructure. Using the H1N1 pandemic experience, we developed a simulation that models the true unobserved influenza incidence trend in the State of Michigan, as well as trends observed at different data collection points of the surveillance system. We calculated the growth rate, or speed at which each trend increases during the pandemic growth phase, and we performed statistical experiments to assess the biases (or differences) between growth rates of unobserved and observed trends. We highlight the following results: 1) emergency-driven high-risk perception increases reporting, which leads to reduction of biases in the growth rates; 2) the best predicted growth rates are those estimated from the trend of specimens submitted to the surveillance point that receives reports from a variety of health care providers; and 3) under several criteria to queue specimens for viral subtyping with limited capacity, the best-performing criterion was to queue first-come, first-serve restricted to specimens with higher hospitalization risk. Under this criterion, the lab released capacity to subtype specimens for each day in the trend, which reduced the growth rate bias the most compared to other queuing criteria. Future research should investigate additional restrictions to the queue.

Keywords

Disease spread modeling, influenza surveillance evaluation, influenza surveillance modeling, respiratory virus surveillance

Respiratory viruses often create emerging outbreaks when unexpected events challenge the immune system of the susceptible population, increase influenza-like illness (ILI) incidence, alter health care–seeking behaviors, and challenge the public health infrastructure created for disease surveillance.[1] In the United States, emerging outbreaks have been either pandemic, such as the 2009 H1N1 and the 2019 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), or seasonal, such as the 2017–2018 A(H3N2). While both pandemic outbreaks resulted from animal-to-human transfer followed by sustained human-to-human transmission, the 2017–2018 A(H3N2) emergency was likely the result of a flu vaccine with less than 25% effectiveness against the predominant A(H3N2) virus.[2]

Although digital surveillance systems have proven beneficial for early warnings of emerging outbreaks (e.g., Pro MED—mail, health maps, and the Global Public Health Intelligence Network),[3] epidemiological
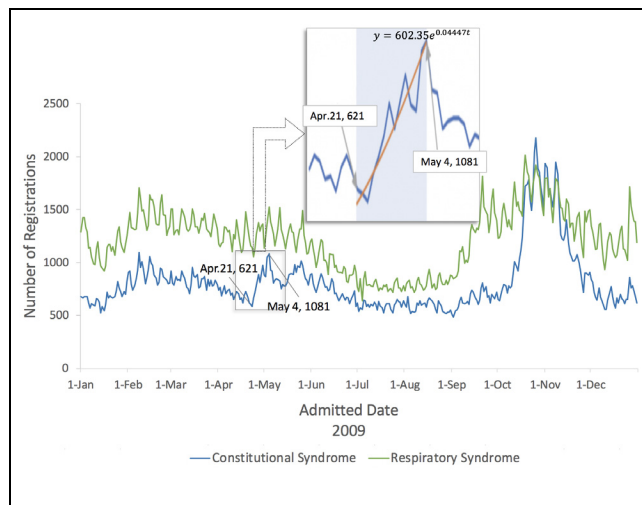
Corresponding Author

Diana Prieto, Johns Hopkins University Carey Business School, 100 International Dr, Rm 1245, Baltimore, MD 21202-4673, USA
(dprieto3@jhu.edu)

characterization of influenza viruses still occurs from official data collection sources such as hospital and emergency department (ED) reports, laboratory testing data, and data from local and state health departments.

As per surveillance standards from the World Health Organization, influenza data collection occurs passively when symptomatic individuals report their symptoms to the health care system.[4,5] In Michigan, for example, EDs are connected to the Michigan Syndromic Surveillance System (MSSS). Registrations of individuals presenting at EDs, including their chief complaint, are automatically and electronically sent into the MSSS, which is monitored by the local and state public health departments. Primary and urgent care practitioners in Michigan can report their weekly ILI cases to the Centers for Disease Control and Prevention's (CDC's) ILI Sentinel Network, but enrollment in the network is optional, and the system does not have the capability for real-time updating of the overall case trend. Aggregate weekly totals and individual case reports from EDs, primary care providers, and urgent care providers, as well as any other type of individual submission (e.g., from schools or flu-testing labs), are captured by the Michigan Disease Surveillance System (MDSS), which is the virtual data-sharing platform across the state and local health departments.

When the influenza pandemic occurred in 2009, it is likely that individuals who sought health care were influenced by the media coverage about the disease. Consider the trend of cases with constitutional and respiratory symptoms that were reported through the MSSS during the nascent phase of the outbreak (Figure 1). Although the respiratory trend does not show a clear pattern, the constitutional trend shows that chief complaints increased after April 21, which is the day when the CDC reported the first human cases of the 2009 H1N1 flu. One might suspect that the increasing trend was influenced by the health care–seeking behavior of individuals who attributed their symptoms to the newly described 2009 H1N1 virus.

Western Michigan University, Kalamazoo, MI, USA (YG, ED); Bronson Methodist Hospital, Portage, MI, USA (RV); University of North Carolina Charlotte College of Health and Human Services, Charlotte, NC, USA (RP); Michigan State University College of Human Medicine, East Lansing, MI, USA (SP); Western Michigan University Homer Stryker MD School of Medicine, Kalamazoo, MI (GS); Johns Hopkins University Carey Business School, Baltimore, MD, USA (DP).

Figure 1 Number of daily emergency department registrations from constitutional symptoms (i.e., fever, headache, malaise, fatigue, and diarrhea) and respiratory symptoms (i.e., shortness of breath or difficulty breathing, cough, sore throat, and runny nose).

Once ILI cases seek health care during routine surveillance, local health departments and sentinel providers send respiratory specimens from a subset of the cases to the state public health laboratory (PHL) or commercial laboratories for further specimen testing and characterization. Local health departments control the submission of specimens to the PHL by following recommended criteria from the Michigan Department of Health and Human Services (MDHHS). In addition, local health departments can receive test results from commercial labs. In April 2009, during the nascent phase of the pandemic, the PHL tested most of the influenza specimens while the MDHHS validated the testing protocols for other laboratories. Also during this phase, the MDHHS recommended that not only sentinels but any health care provider could send specimens to the PHL. This operational landscape, together with the lab-testing capacity and the manual methods for receiving and processing the specimens, created congestion and delays of up to 2 weeks.[6–8] Once the PHL became congested, the MDHHS restricted the specimen submission to prioritize specific populations with severe ILI presentations (e.g., hospitalized patients and pregnant women). This reactive response could have created biases (or differences) between the incidence trends observed through surveillance and the real unobserved trends.

Significant biases may produce underreaction or overreaction in mitigation efforts. As an example, consider the 2009 and the 1918 influenza pandemic outbreaks.

The 2009 pandemic had a median growth rate of 0.125 while the 1918 pandemic had a median growth rate of 0.201 (both rates were calculated from data in Biggerstaff et al.[9]). If the true unobserved outbreak is like the 2009 outbreak but the data show a 1918 pandemic-like outbreak, the bias of –0.076 may lead authorities to implement aggressive travel-related control policies, and overreaction may occur. Conversely, if the true unobserved outbreak is like the 1918 outbreak but data show a 2009 pandemic-like outbreak, the bias of 0.076 would trigger underreaction.

In this article, we present a simulation-based evaluation of surveillance operations for reporting, collection, and testing of specimens during the nascent phase of an emerging influenza outbreak in the State of Michigan. We seek to accomplish 2 objectives: the first is to identify factors that contribute to differences between real and observed incidence trends. Table 1 describes the complete pool of factors to explore, which include "pandemic risk perception," "primary care access," and "sampling criteria" to prioritize specimens in the PHL. The second objective is to quantify how well the observed trends represent the behavior of the true incidence trend. To our knowledge, these 2 objectives have not been investigated in previous simulation models for surveillance research. Simulation models have been used to explore policies for spatial allocation of data providers[10,11] and to test the performance of process monitoring algorithms for detecting outbreaks of influenza[12] or other diseases.[13–15]

Our simulation model is built as a realistic representation of cocirculating influenza viruses across multiple regions in the State of Michigan, in the presence of mitigation, containment, and surveillance policies. Epidemiological and surveillance processes are grounded on real data, public health documentation, and field experts in public health operations who were in service during the 2009 influenza pandemic.

Methods

Model

We developed a simulation that models the true unobserved influenza incidence trend in the State of Michigan, as well as trends observed at different data collection points of the surveillance system. Figure 2 presents the different stages at which the trends are generated.

Stage 0 generates the true unobserved influenza incidence trend (Figure 2). In this stage, 2 cocirculating influenza viruses spread through a network of regions simulating human contact patterns in the State of Michigan (~10,000,000 inhabitants). The Michigan network connects its 4 characteristics regions: Upper Peninsula, Northern Lower Peninsula, Southeast Michigan Council of Governments, and Southern Lower Peninsula.

Real data were used to build the geographic and demographic features, as well as the contact and travel patterns within and between the regions and across the Michigan boundaries. Data sets were sourced from the US Census Bureau,[16–18] National Household Travel Survey,[19] Michigan School Data webpage,[20] National Center for Education Statistics,[21] National Institute for Early Education Research,[22] Michigan Department of Transportation,[23–25] and survey data[26] (see Supplementary [SI]: 1.1–1.3).

The 2 cocirculating viruses were seeded to re-create an emerging virus that interacts with a seasonal strain already infecting the population. The emerging and seasonal viruses were simulated using the epidemiological features of the pandemic and seasonal H1N1 strains, respectively.[27] Cocirculation enables the further sampling and testing of influenza viruses in the surveillance system. The viruses are seeded in the Southern Lower Peninsula and spread through contacts made by infected individuals while traveling within and between regions. Infected individuals are also allowed to enter and leave Michigan by traveling across the state boundaries.

Antiviral and seasonal vaccination have been modeled as factors that reduce the probability of infecting a susceptible individual, using existing estimates of influenza antiviral and vaccine efficacy (see SI: 1.4). Some sick individuals withdraw from their usual schedules and reduce the population of infectious individuals in the community. A fraction of these individuals will be hospitalized based on age-based hospitalization rates in Michigan (see SI: 1.4).

The previous contact and mitigation structure generates the daily trend of original infected cases with realistic variability and provides a robust simulation testbed for a wide variety of scenarios.

In addition to stage 0, 3 more stages were simulated to represent the surveillance system as shown in Figure 2 (see SI: 1.5). In stage 1, symptomatic individuals seek health care, based on a set of disease and behavioral factors that have been previously established in the literature as influencing health care–seeking behavior and introducing delay in the disease reporting (Table 1, stage 1). Stage 1 generates the trend of daily constitutional cases that are reported to the MSSS (MSSS trend). In stage 2, health care providers collect and/or submit influenza specimens to the PHLs, based on a set of screening factors (Table 1, stage 2), which generates the trend of specimens that are submitted to the PHLs. In stage 3, cases are tested based on a set of operational factors,

Table 1 Experimental Factors Considered in the 3 Stages[a]

Stage	Factor Name	Description	Factor Type	Values during Experimentation
Stage 1: Cases seeking health care and reporting to MSSS	Severity of symptoms	Are the symptoms of the individual severe enough to seek health care?	N	$0 - 1^{R+}$
	Perception to be pandemic	Probability that an individual perceives his or her symptoms to be pandemic	C	$0 - 1^{R+}$
	Health care schedule	Whether the patient is seeking health care on a weekend or not	N	0—weekend 1—not weekend
	Primary care access	Probability that an individual has access to primary care	C	$0 - 1^{R+}$
	ED v. urgent care	Probability that an individual goes to urgent care instead of the ED. Higher probability indicates that more patients are going to urgent care.	C	$0 - 1^{R+}$
	MSSS reporting	Probability that the symptoms of a patient are reported in the MSSS	F	0.7
Stage 2: Specimen collection and submission by health care providers	Collect sample	Probability that a sample is reported by physicians or urgent care providers	C	$0 - 1^{R+}$
	Shipping schedule	Whether the shipping of a specimen occurs during a weekend or not	N	0—weekend 1—not weekend
	Remote shipping	Whether the shipping of a specimen occurs from a remote area or not	N	0—no 1—yes
Stage 3: Specimen testing in the PHL	PHL schedule	Whether the PHL works on weekends or not	C	0—weekend 1—not weekend
	Sampling criteria	Sampling criteria for prioritizing the testing of specimens in the PHLs	C	1—FCFS 2—FCFS restricted to high-risk groups (restricted FCFS) 3—switch from FCFS to restricted FCFS when testing capacity limit is reached 4—use FCFS for the first week of pandemic declaration and stop testing afterward 5—use restricted FCFS for the first week after pandemic declaration and stop testing
	Specimen longevity	Whether a specimen has been in PHL for more than 3 days or not	N	0—no 1—yes
	PHL capacity	Testing capacity limit of the PHL per day	C	$100 - 1,000^{I+}$

C, controllable; ED, emergency department; F, fixed; FCFS, first come, first serve; N, noncontrollable; MSSS, Michigan Syndromic Surveillance System; PHL, public health laboratory.

[a]R+: the factor value is a positive real number; I+: the factor value is a positive integer.
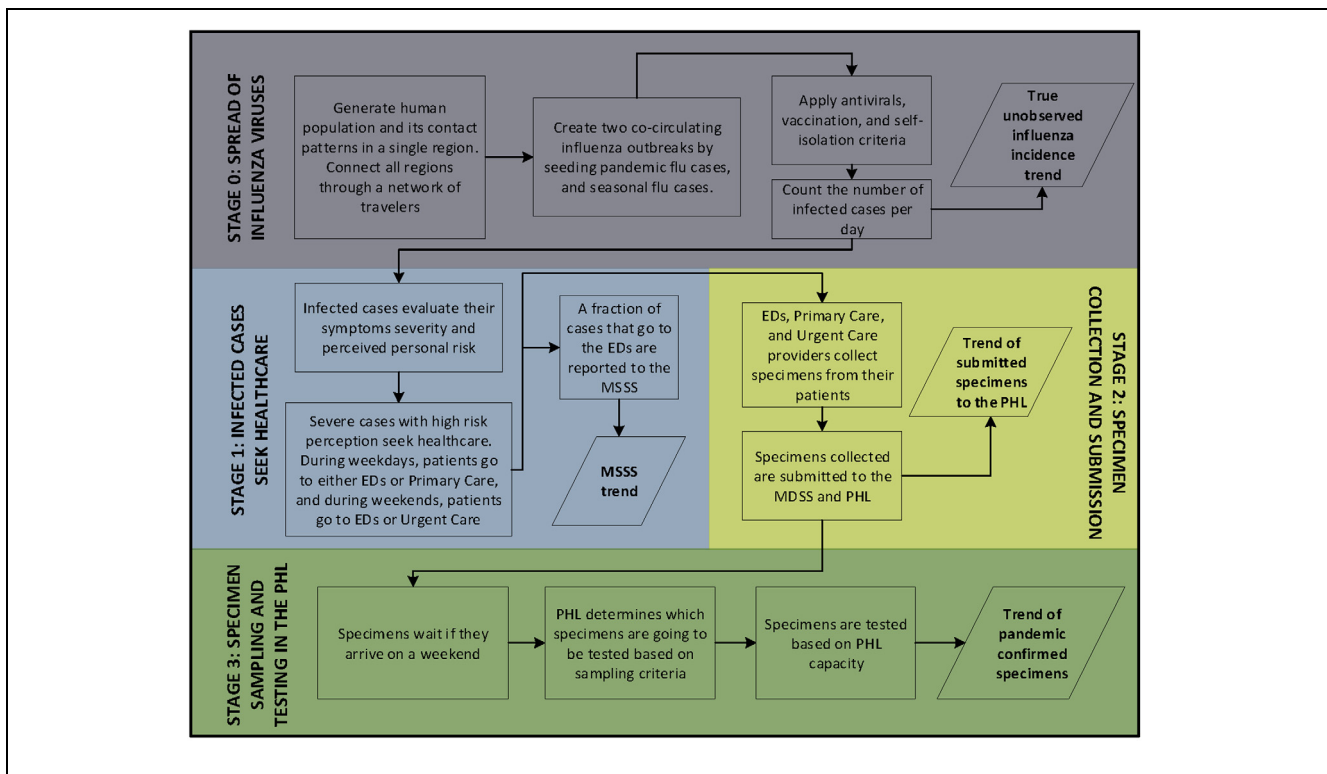
Figure 2 Simulation of influenza incidence trends. Stage 0 generates the trend of infected cases per day, which is not observed by the surveillance system. Stages 1, 2, and 3 generate incidence trends at different surveillance points. MSSS, Michigan Syndromic Surveillance System; PHL, public health laboratory.

including PHL schedule, sampling criteria, and PHL capacity. Factor experimental values were selected based on observed operational conditions during the 2009 pandemic influenza (Table 1, stage 3). Stage 3 generates the trend of pandemic confirmed specimens by date of submission to the PHL. Figure 3 shows 2 scenarios that present the simulated surveillance trends for the growth phase of an influenza outbreak.

Experiments

Two experiments were conducted for calibration, validation, sensitivity, and scenario analysis. The first experiment sampled from stage 0 of the model, with the objective to calibrate the model using the simulated ILI growth rate r_0 and validate the parameter settings resulting from the calibration. For the calibration, the growth rate r_0 was fitted to the real growth rate of the pandemic outbreak in the State of Michigan. The real growth rate was extracted from the MSSS trend of constitutional case daily incidence (see Figure 1). The growth rate was

estimated by fitting an exponential growth equation to each of the trends (see SI: 2.2).

To fit r_0, the experiment sampled a selected set of scenarios that allowed the creation of 2 second-order polynomial equations, where the response variables were r_0 and η (standard error of the r_0 estimate), respectively, and the predictors of both equations were the model parameters in Table 2. The second-order equations provided the linear, interaction, and quadratic terms to characterize the surfaces of r_0 and η. Using optimization and gradient search algorithms, 2 scenarios were found that minimize the difference between the r_0 and η parameters modeled with the second-order equations and the real values for these 2 parameters from the MSSS data. The calibration methods are explained in detail in Section 2 of the supplementary file (see SI: 2).

Table 3 shows the 2 scenarios obtained. Face validation was conducted by discussing the scenarios with the public health operations experts on our team. It was concluded that both hypothetical scenarios portray plausible initial conditions of the pandemic outbreak in Michigan,
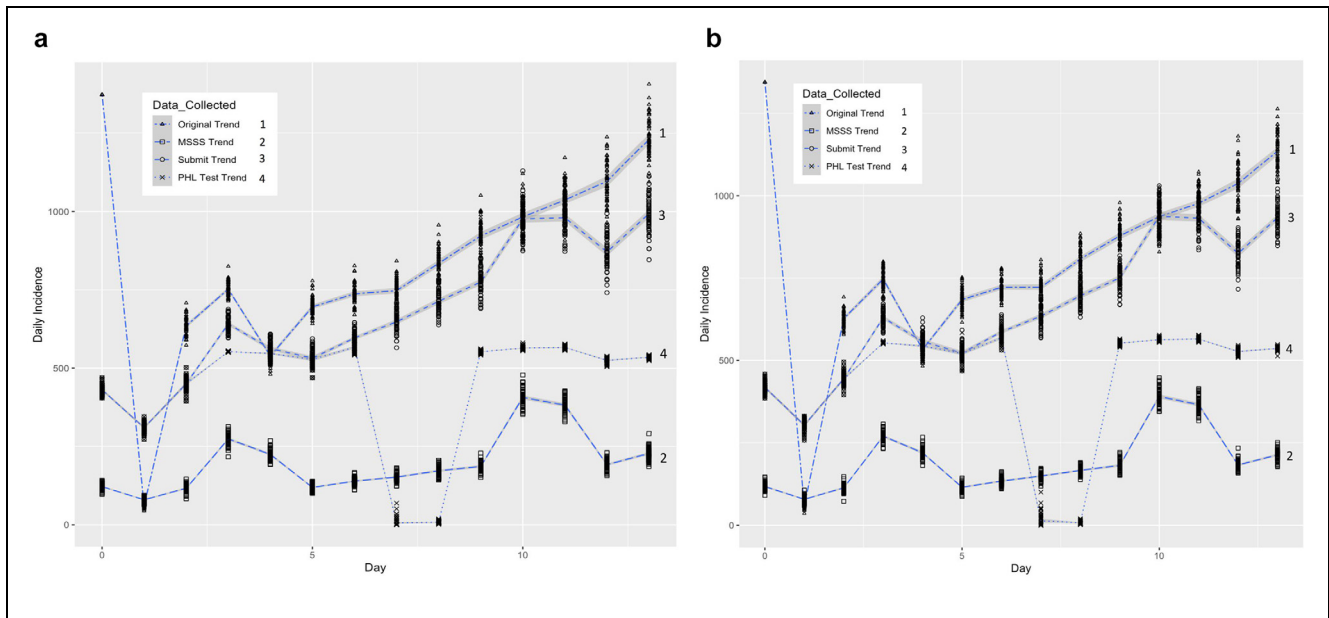
Figure 3 (a) Scenario 1: simulated data collection trends from 60 replicates using expected reproduction numbers of 2.0965 and 1.6 for the pandemic and seasonal outbreaks, respectively; 686 initial infected cases for the pandemic outbreak and 686 initial infected cases for the seasonal outbreak; and a value of 1 (high) disease severity and a 26% chance that a vulnerable person receives antiviral treatment. (b) Scenario 2: simulated data collection trends from 60 replicates using expected reproduction numbers of 1.96 and 1.6 for the pandemic and seasonal outbreaks, respectively; 658 initial infected cases for the pandemic outbreak and 686 initial infected cases for the seasonal outbreak; and a value of 1 (high) disease severity and a 0.001% chance that a vulnerable person receives antiviral treatment. Both scenarios have the following experimental parameters in the surveillance system: perception to be pandemic = 0.5, primary care access = 0.5, emergency department (ED) v. urgent care = 0.5, collect sample = 0.5, public health laboratory (PHL) capacity = 515, PHL schedule = 0, and sampling criteria = 1 (first come, first serve). Both scenarios are run to collect 14 days of the cocirculating outbreak.

Table 2 Calibration Parameters

Parameter	Description	Parameter Type	Experimental Values and Ranges
$E[R_1]$	Expected reproduction number—pandemic	C	$1.6–2.1^{R+}$
$E[R_2]$	Expected reproduction number—seasonal	C	$1.1–1.6^{R+}$
I_1	Initial number of infected cases—pandemic	C	$30–1,000^{I+}$
I_2	Initial number of infected cases—seasonal	C	$30–1,000^{I+}$
σ	Probability of withdrawal from regular activities	C	$0.00001–0.99999^{R+}$
a	Proportion of risk group population receiving antivirals	C	$0.00001–0.99999^{R+}$
s_2	Whether seasonal vaccination is in place or not	F[a]	1 (in place)
q	Whether self-induced absenteeism and hospitalization were in place or not	F[a]	2 (In place)

C, controllable; F, fixed.

[a]Parameters s_2 and q were fixed since routine vaccination, self-induced absenteeism, and hospitalization were already in place before pandemic declaration.

assuming the overall case incidence grows similarly to the MSSS trend of constitutional cases: high pandemic reproduction number, moderate seasonal reproduction number, high withdrawal percentage, and low antiviral usage coverage. In addition, the initial pool of infected

cases was estimated to be high (686 in both scenarios), which suggests that influenza spread initiated before the pandemic was declared.

The second experiment sampled from all the stages in the simulation model, with the objective to provide

Table 3 Optimal Calibrated Scenarios

$E[R_1]$	$E[R_2]$	I_1	I_2	σ	a	$\hat{r}_0$ (CI)	$SE[\hat{r}_0]$ (CI)
2.0965	1.6	686	686	0.99999	0.26	0.0449 (0.0437, 0.046)	0.00232 (0.00231, 0.00233)
1.9606	1.6	658	686	0.99999	0.00001	0.0402 (0.039, 0.0414)	0.00234 (0.00233, 0.00235)

CI, confidence interval.

sensitivity analysis and external validation. In addition, the experiment sought to provide understanding of the explanatory power of the incidence trends at different stages of the surveillance system. A set of scenarios was sampled to create second-order polynomial equations, where the responses were the biases between the growth rates of the trends generated by the surveillance system. The biases can be defined as follows:

$$\delta_i = r_0 - r_i, \forall i \neq 0, \qquad (1)$$

$$\delta_{i,1} = r_{0,1} - r_{i,1}, \forall i \neq 0, \qquad (2)$$

where r_i is defined as the ILI growth rate of the trends simulated in surveillance stage i, where $i = 0, 1, 2, 3$. Also, $r_{i,1}$ is defined as the simulated growth rate in stage i for the incidence trend of the pandemic-like virus. Note that $\delta_{3,1}$ was the only response modeled out of the biases in the pandemic trends, as $r_{3,1}$ is the only observed pandemic trend. For all the equations, the predictors were the surveillance controllable factors in Table 1. In the multivariate sensitivity analysis, analysis of variance (ANOVA) was conducted on the equation coefficients to determine the significant effects of linear, interaction, and quadratic terms in δ_i and $\delta_{3,1}$ (see SI: 2.3).

For external validity, we compared real and simulated growth rates of pandemic cases confirmed by PHL testing. MDHHS 2009 data on pandemic confirmed cases were used to estimate a real growth rate from April 19, 2009, to May 30, 2009 (week 16 to week 21).[28,29] To determine whether this real growth rate could be produced by the simulation, experimental runs that used first come, first serve (FCFS) as the sampling criteria were selected for the analysis. This selection considered that the real sampling in the PHLs was conducted mostly using FCFS. From the growth rates of the selected runs, a confidence interval was calculated and compared against the real growth rate of pandemic cases. Since the confidence interval includes the real growth rate from confirmed cases, we confirm the validity of our model to replicate this real scenario.

The second experiment generated 3,200 replicates, consisting of 320 combinations with 10 replicates. For each of the 320 combinations, we estimated and compared the fraction of the variability explained by the predictors in each of the following relationships:

$$r_0 = \phi_i r_i, \forall i \neq 0, \qquad (3)$$

$$r_{0,1} = \phi_{i,1} r_{i,1}, \forall i \neq 0, \qquad (4)$$

where ϕ_i and $\phi_{i,1}$ are constants. We then counted the number of combinations where each r_i and $r_{i,1}$ explained the highest levels of variability.

Results

Figure 4 shows the results for δ_1 in the first optimal scenario, which can be interpreted as the bias between the true unobserved influenza incidence trend and the MSSS trend. "ED v. Urgent Care" and the interaction between the "perception to be pandemic" and "primary care access" were found to be significant. The U shape in the first graph of Figure 4 results from the selection of health care provider based on the day of the week: during weekdays, patients go to either EDs or primary care, and during weekends, patients go to EDs or urgent care (see Figure 2 and SI: 1.5). When the "ED v. urgent care" probability is less than 0.25, less patients are routed to urgent care and more to the EDs during both weekdays and weekends, which reduces δ_1. There is an indifference zone between 0.25 and 0.75 where removing patients from the EDs does not alter the bias. Probability values higher than 0.75 increase the bias, since during weekends, most of the patients go to urgent care and not to EDs, and the lack of samples in these 2 days reduces the capability of the MSSS growth rate to represent changes in the incidence. In the second graph of Figure 4, primary care access higher than 25% implies that less patients are reported to the ED, which increases δ_1 at any level of pandemic perception. Similar results are obtained for the second optimal scenario (which are available in SI: 3.4.1).

Figure 5 shows the results for δ_2, which can be interpreted as the bias between the true unobserved influenza incidence trend and the trend of specimens submitted to
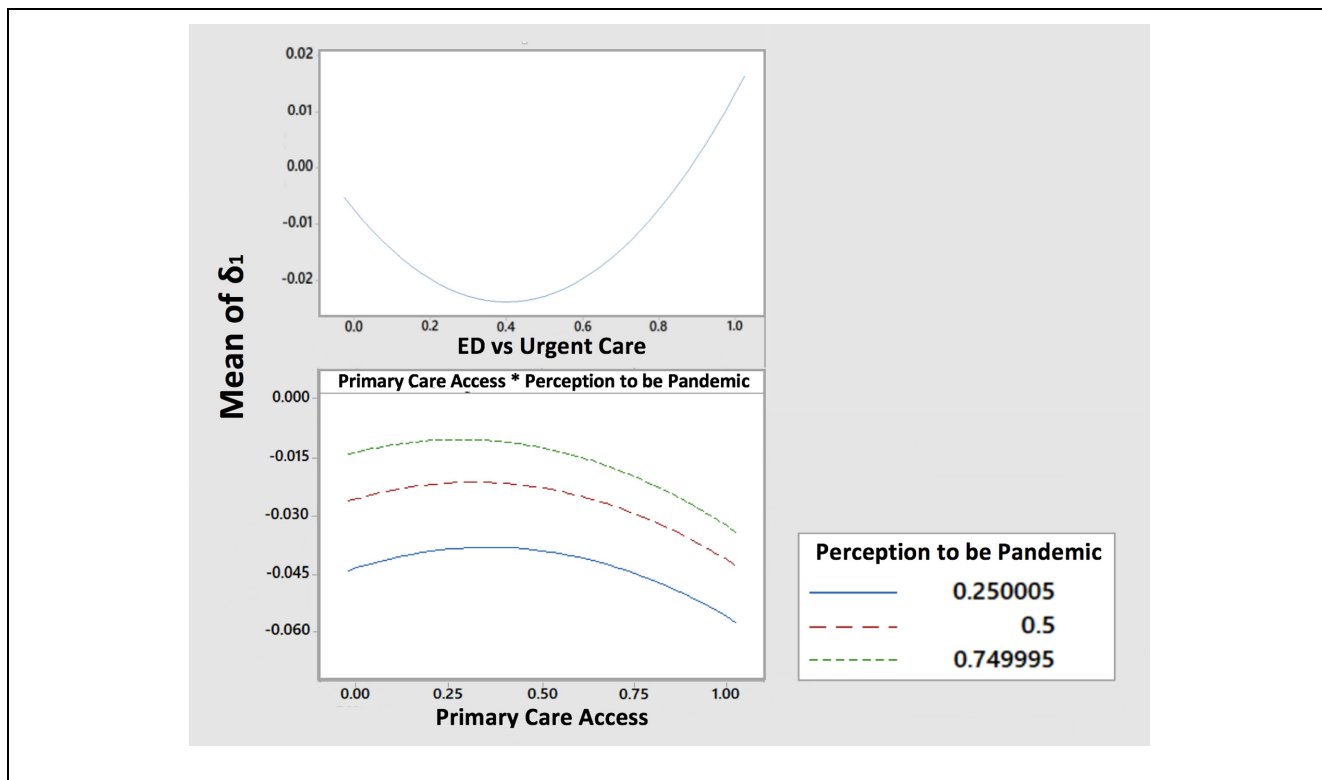
Figure 4 Effect plot of factors affecting bias between original and Michigan Syndromic Surveillance System trend. ED, emergency department.

the PHL. The first column shows that a high "ED v. urgent care" probability (≥ 0.75) leads to a higher difference in δ_2 since cases arriving during the weekend are lost when health care providers do not report them (i.e., when the probability to collect samples is low), and this effect is observed across all the interactions. The second column shows that higher "perception to be pandemic" routes more patients with severe symptoms to the PHL, which reduces δ_2 across all the interactions. In the third column, higher probability to collect samples from primary care and urgent care also routes more patients to the PHL, thereby reducing δ_2 across all the interactions. Column 4 shows a slightly decreasing effect in the bias as the "primary care access" increases. From row 3 in Figure 5, note that when the "collect sample" probability is low, the effect of "primary care access" is lower than the effect of "ED v. urgent care" in the bias, since high primary care access contributes to enough cases submitted during weekdays, which better shape the trend of PHL submission compared to the few cases submitted during weekends contributed by urgent care. Similar results are obtained for the second optimal scenario (see SI: 3.4.2).

Figure 6 shows the results for $\delta_{3,1}$ in the first optimal scenario, which translates as the bias between the true unobserved pandemic trend, and the trend of specimens that tested positive for pandemic flu once submitted to the PHL. Note that FCFS and FCFS restricted to high-risk groups always yield the lowest biases across all the interactions, which indicates that switching the sampling criteria while the outbreak is progressing (as occurs with the other 3 strategies tested) is detrimental for the estimation of $r_{3,1}$. "Perception to be pandemic," "primary care access," "ED v. urgent care," and "collect sample" present similar behavior as in the previous stages across all the interactions. Figure 6e shows that whether the lab works on weekends or not does not affect the bias under FCFS or FCFS restricted to high-risk groups. Figure 6f shows that under FCFS and low PHL capacity, the bias is initially high as the lab cannot test enough specimens that represent the end of the reporting period. The bias gets close to zero when the capacity is around the 550 specimens. However, after the 550 specimens, there is enough capacity to timely test all the specimens in the queue and observe the bias associated with testing all the samples that arrive at the PHL. Under FCFS restricted
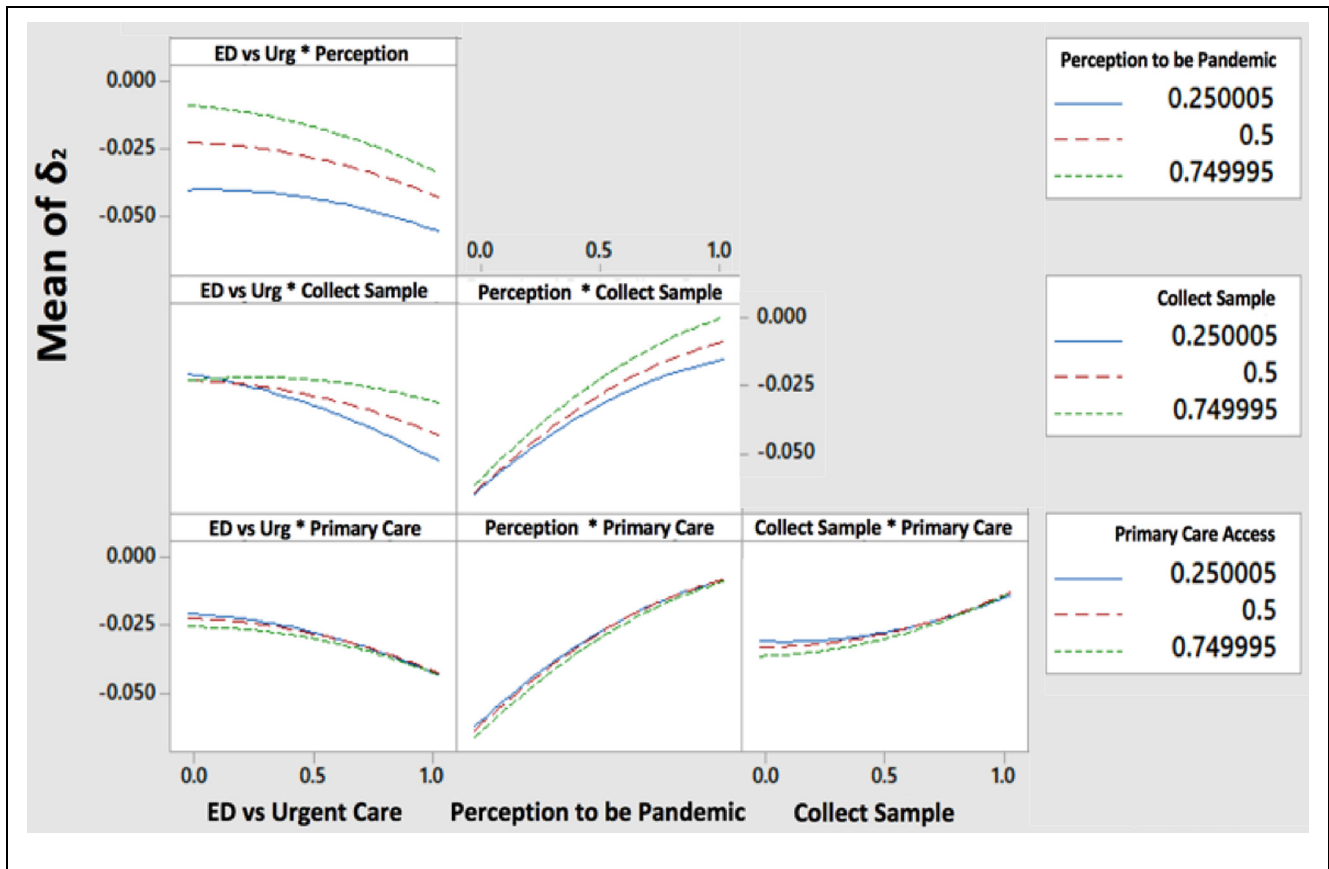
Figure 5 Effect plot of factors affecting bias between original and submit trend. ED, emergency department.

and low PHL capacity, the restriction reduces the queue length, and it is possible to timely test enough specimens to characterize the real pandemic growth rate. However, after the 325 specimens, there is enough capacity to timely test the queue and observe the bias associated with testing all the samples that arrive at the PHL. These results indicate that it is possible to minimize the bias while controlling the PHL capacity. The results and interpretation of δ_3 are similar to those of $\delta_{3,1}$ (see SI: 3.4.4).

Explanatory Power of the Trends

Table 4 and Figure 7 show that r_2, which is the growth rate of all specimens submitted to the PHL, is the rate that better explains the true ILI rate in 309 of 320 cases. In addition, r_2 is the rate that better explains the true pandemic trend in 304 of 320 cases. Similar results were obtained for $r_{i,1}$. The results shown here are for the first optimal scenario, but similar results were also obtained for the second optimal scenario (see SI: 3.4.8).

Table 4 Number of Combinations Explained by the Observed Growth Rates under the First Optimal Scenario

Growth Rate	r_1	r_2	r_3
Overall	7	309	4
Pandemic	13	304	3

Discussion

Most of the surveillance system can be seen as a set of connected pipelines through which cases are constantly routed. The parameters "perception to be pandemic," "primary care access," "ED v. urgent care," and "collect sample" act as valves that divert or route the flow of specimens to the data collection points. In general, opening a valve that leads to a data collection point contributes to cases that reduce the bias, and closing a valve has the opposite effect.

"Perception to be pandemic" is one of the first valves of our simulated surveillance system and can route or
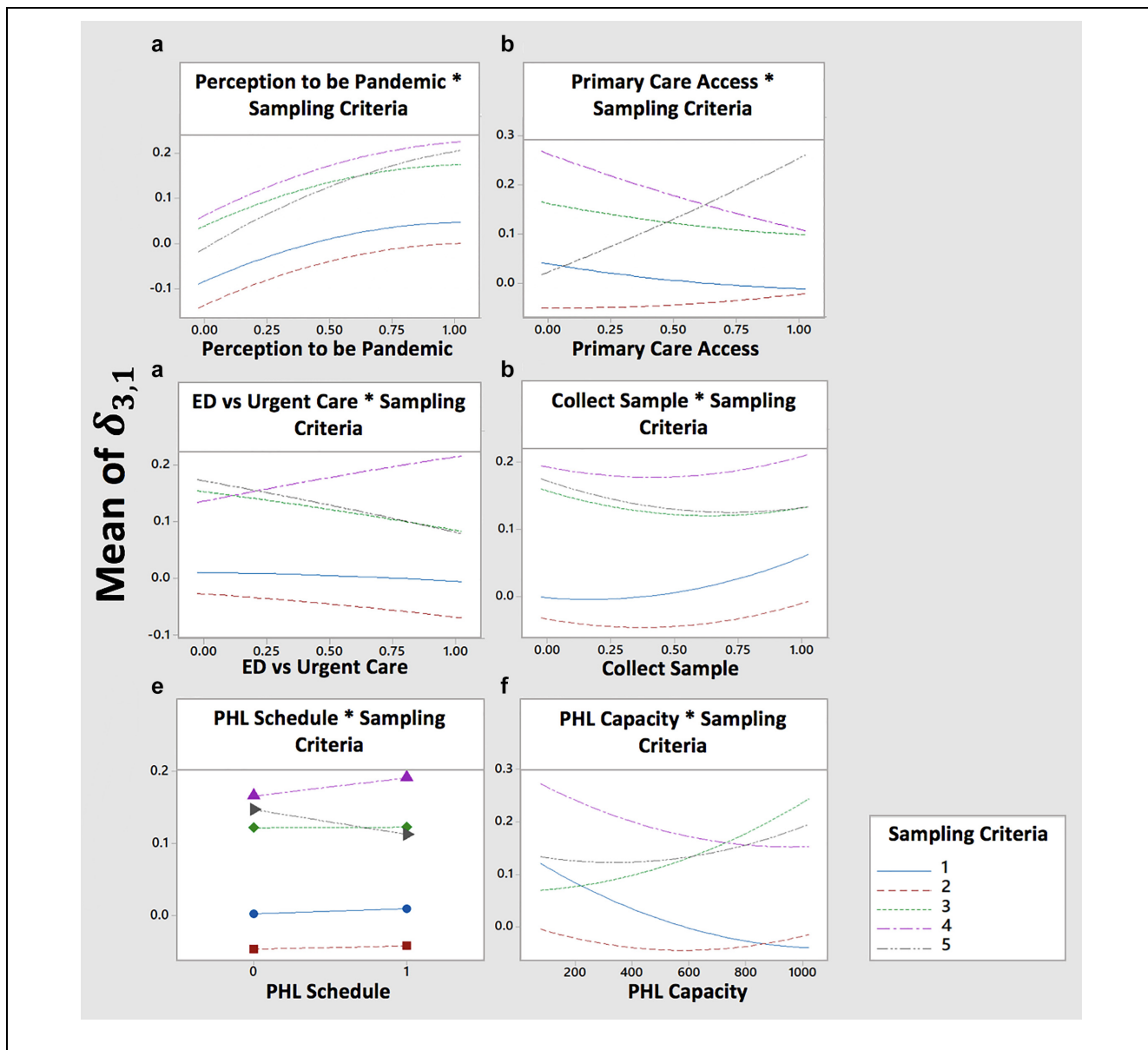
Figure 6 Effect plot of factors affecting bias between original and public health laboratory (PHL) tested trend. ED, emergency department.

divert a high volume of cases. Hence, it has a highly significant impact on all the biases under study. Our results suggest that high risk perception due to the pandemic increases the reporting, which leads to the reduction of biases in the growth rates. As the public interest decreases, so will the reporting pattern.[30] Therefore, an emerging outbreak should be seen as an opportunity to support the collection of larger sample sizes, and public communication should provide a realistic sense of the risk posed by the influenza virus.

"ED v. urgent care" is a key valve in routing or diverting cases reporting over the weekend, and hence it is important to encourage urgent care and EDs to report their weekend cases. This practice will likely produce more representative samples in the incidence trends for each of the weekend days reported. The PHL should also consider that misrepresentation may occur for days with small sample sizes.

Once the specimens arrive at the PHL, the bias from the true unobserved pandemic curve is affected by 2
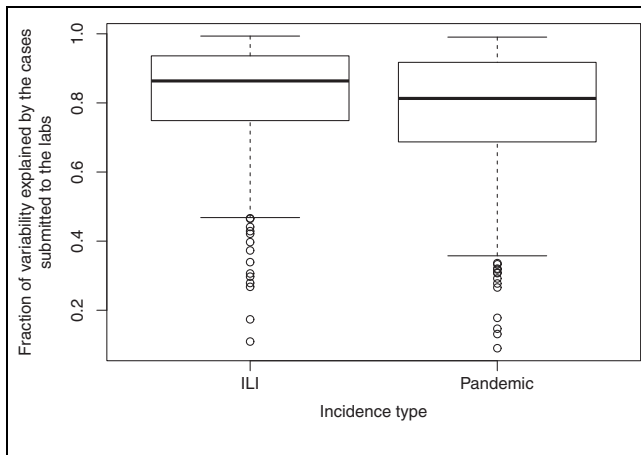
Figure 7 Boxplots of the variability in the overall ILI and pandemic growth rates that is explained by the growth rates of cases submitted to the lab.

factors: "sampling criteria" and "PHL capacity." Based on the sampling decisions made in Michigan during the H1N1 pandemic, we designed and tested 5 types of "sampling criteria." Two of these criteria implied that the tested sampling method was maintained over the course of the growth phase, while 3 of these criteria implied a switch in the sampling method. Our results show that switching from one sampling criterion to another contributes to higher biases than maintaining the same sampling criteria. When the transition happens from FCFS or FCFS restricted to no sampling, the growth rate is estimated only with 1 week of data, which increases the bias. When the transition happens from FCFS to FCFS restricted, the pool of samples is reduced, and therefore the samples that tested positive out of this pool do not represent incidence at the same scale as the FCFS samples. Estimation methods should ensure that proper adjustment strategies are used to account for abrupt changes in the pool of samples.

Under constrained capacity, FCFS restricted outperformed the other 4 sampling criteria when the objective was to minimize the growth rate bias for the overall incidence trend of pandemic cases. The restriction allowed that the PHL tested cases for each day that would have otherwise expired if they had continued waiting in the queue. Hence, the restriction releases capacity to collect samples for each day in the trend, which reduces the growth rate bias. Future research should aim at investigating additional restriction strategies that further minimize the bias.

Although our results were obtained using the H1N1 pandemic scenario, we believe that our insights can be considered during other emergent outbreaks of respiratory viruses. During the H1N1 pandemic, commercial labs were still working on the approval of their testing protocols and not reporting during the first 2 weeks after pandemic declaration. During the 2019 SARS-CoV-2, it took more than a week and a half after pandemic declaration to begin operationalizing government-funded free testing across the United States. These experiences illustrate that during the nascent phase of an outbreak emergency, the PHL is at the frontline of outbreak detection and follow-up. However, their testing capabilities are usually constrained. Our results suggest that, under conditions of high specimen reporting and submission, incidence growth is better characterized in real time if the PHL provides the results from a limited sample representing each day of the curve, rather than providing results from all the arriving specimens. In practice, the PHL labs are testing not only to characterize the epidemic curve but also to detect changes in the virus genetic makeup, as well as to identify positive cases guiding targeted mitigation (e.g., isolation of cases and quarantine of case contacts). Although the PHL capacity planning cannot be driven only by the epidemiological characterization, our results suggest that the PHL could report the results in 2 categories: 1) the results of a full set of specimens tested to comply with the different objectives and 2) out of the results from the full set, a fraction specifically selected for epidemiological characterization. In our research, the daily fraction was dictated by the specimens arriving at the PHL that met our age-based risk group definition. This definition worked well to represent the population since in our model, the disease was affecting the population homogeneously across age groups. In practice, the fraction to consider should be detached from categories representing heterogeneous incidence.

Our results also show that the growth rate of ILI cases submitted to the PHL has a good explanatory power for the growth rate of overall ILI cases. Similarly, the growth rate of pandemic cases submitted to the PHL has a good explanatory power for the growth rate of overall pandemic cases. This result seems to illustrate that the node connected to the most and the widest variety of health care providers in the surveillance network also collects the most representative data, leading to lower biases in the growth rates. In countries like England, public health authorities are in search of better predictors of ILI activity, to complement or replace the existing primary care consultation data.[30] Convergent data sources that receive

reporting from different types of providers (i.e., urgent care, primary care, or hospital) might present as enhanced alternatives to primary care provider consultation for the estimation of ILI incidence. In Michigan, the MDHHS receives most of the requests for PHL testing before specimens are being sent, but some are submitted directly to the PHL during an emergency. The PHL sometimes declines specimens directly reported due to capacity constraints, and therefore, there are no results to submit to the MDSS. Hence, the MDSS would produce the least biased growth rates, as long as the PHL underreporting is considered (e.g., by means of sensitivity analysis).

To strengthen viral surveillance, several future research directions can be proposed for the modeling framework that we have established:

- **Generation of synthetic data**. Our framework generates incidence trends at different points of data collection. Such trends can be helpful to validate and compare the performance of different estimation and forecasting methods.
- **Simulation-based learning for operational decision making**. Our framework can be used in tabletop exercises by allowing decision makers to play and learn about disease progression from different simulated scenarios. It would be valuable to understand whether emergency planning is enhanced by such learning strategies.
- **Better distribution of the PHL testing resources**. Novel sampling schemes that aim to better represent incidence across day, age, and geography can minimize the biases while constrained by the PHL testing capacity. Development and testing of these schemes could provide implementable recommendations.

ORCID iDs

Yuwen Gu https://orcid.org/0000-0002-0862-7276
Diana Prieto https://orcid.org/0000-0003-4939-0786

Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making* website at http://journals.sagepub.com/home/mdm.

References

1. Briand S, Mounts A, Chamberland M. Challenges of global surveillance during an influenza pandemic. *Public Health*. 2011;125(5):247–56.
2. Flannery B, Chung JR, Belongia EA, et al. Interim estimates of 2017–18 seasonal influenza vaccine effectiveness—United States, February 2018. *MMWR Morbid Mortal Wkly Rep*. 2018;67(6):180–5.
3. Magid A, Gesser-Edelsburg A, Green MS. The role of informal digital surveillance systems before, during and after infectious disease outbreaks: a critical analysis. In: Radosavljevic V, Banjari I, Belojevic G, eds. *Defence against Bioterrorism*. Dordrecht: Springer Netherlands; 2018. p 189–201.
4. World Health Organization. WHO global technical consultation: global standards and tools for influenza surveillance, March 2011.
5. Morse SS. Global infectious disease surveillance and health intelligence. *Health Affairs*. 2007;26(4):1069–77.
6. Prieto D, Das T, Savachkin A, Uribe A, Izurieta R, Malavade S. A systematic review to identify areas of enhancements of pandemic simulation models for operational use at provincial and local levels. *BMC Public Health*. 2012;12(251).
7. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comp Biol*. 2015;11(10):e1004513.
8. Santillana M, Nguyen AT, Louie T, et al. Cloud-based electronic health records for real-time, region-specific influenza surveillance. *Sci Rep*. 2016;6:25732.
9. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Public Health*. 2014;14(480).
10. Souty C, Boëlle P-Y. Improving incidence estimation in practice-based sentinel surveillance networks using spatial variation in general practitioner density. *BMC Med Res Methodol*. 2016;16(1):156.
11. Scarpino SV, Dimitrov NB, Meyers LA. Optimizing provider recruitment for influenza surveillance networks. *PLoS Comp Biol*. 2012;8(4):e1002472.
12. Cao P-H, Wang X, Fang S-S, et al. Forecasting influenza epidemics from multi-stream surveillance data in a subtropical city of China. *PLoS One*. 2014;9(3):e92945.
13. Lopes Antunes AC, Jensen D, Halasa T, Toft N. A simulation study to evaluate the performance of five statistical monitoring methods when applied to different time-series components in the context of control programs for endemic diseases. *PLoS One*. 2017;12(3):e0173099.
14. Dórea FC, McEwen BJ, McNab WB, Sanchez J, Revie CW. Syndromic surveillance using veterinary laboratory data: algorithm combination and customization of alerts. *PLoS One*. 2013;8(12):e82183.
15. Dupuy C, Morignat E, Dorea F, Ducrot C, Calavas D, Gay E. Pilot simulation study using meat inspection data for syndromic surveillance: use of whole carcass

condemnation of adult cattle to assess the performance of several algorithms for outbreak detection. *Epidemiol Infect.* 2015;143(12):2559–69.

16. US Census Bureau. America's families and living arrangements. 2010. Available from: https://www.census.gov/population/www/socdemo/hh-fam/cps2010.html

17. US Census Bureau. Profile of general population and housing characteristics: 2010 demographic profile data. 2010. Available from: https://factfinder.census.gov/faces/tableservices/jsf/pages/ productview.xhtml?src = CF

18. US Census Bureau. Statistics of U.S. businesses (SUSB). 2010. Available from: Retrieved from https://www.census.gov/programs-surveys/susb.html

19. National Household Travel Survey (NHTS). 2009 datasets. 2009. Available from: https://nhts.ornl.gov/downloads

20. MI School Data. K–12 school data file. 2010. Available from: https://www.mischooldata.org/DistrictSchoolProfiles2/EntitySummary/SchoolDataFile.aspx

21. National Center for Education Statistics (NCES). Michigan. 2013. Available from: https://nces.ed.gov/collegenavigator/?s = MI

22. National Institute for Early Education Research (NIEER). Michigan profile [data file]. 2013. Available from: http://nieer.org/wp-content/uploads/2014/10/Michigan_2013-1.pdf

23. Comprehensive Household Travel Data Collection Program. Available from: http://www.michigan.gov/documents/mdot/mdot_MTC_II_comparison_report_333906_7.pdf

24. Travel characteristics technical report. Michigan Department of Transportation. Available from: https://www.michigan.gov/documents/MDOT_TravCharTR_Final20060804_167340_7.pdf

25. Michigan Department of Transportation. State long range transportation plan 2005–2030: Travel characteristics technical report. 2006. Available from: https://www.michigan.gov/documents/MDOT_TravCharTR_Final20060804_167340_7.pdf

26. Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 2008;5(3):381–91.

27. Prieto D, Das TK. An operational epidemiological model for calibrating agent-based simulations of pandemic influenza outbreaks. *Health Care Manage Sci.* 2016;19(1): 1–19.

28. Michigan influenza surveillance summary 2008–2009 influenza season. 2009. Available from: https://www.michigan.gov/documents/mdch/2008-2009_Influenza_Season_Summary_331956_7.pdf

29. Michigan influenza surveillance summary 2009–2010 influenza season. 2010. Available from: https://www.michigan.gov/documents/mdch/2009-2010_Influenza_Season_Summary_331957_7.pdf

30. Birrell PJ, Pebody RG, Charlett A, Zhang X-S, De Angelis D. Real-time modelling of a pandemic influenza outbreak. *Health Technol Assess.* 2017;21(58):1.