



Estimating the Effectiveness of Early Control Measures through School Absenteeism Surveillance in Observed Outbreaks at Rural Schools in Hubei, China

Yunzhou Fan¹, Mei Yang², Hongbo Jiang¹, Ying Wang¹, Wenwen Yang¹, Zhixia Zhang¹, Weirong Yan^{1,3}, Vinod K. Diwan³, Biao Xu⁴, Hengjin Dong⁵, Lars Palm⁶, Li Liu^{1*}, Shaofa Nie^{1*}

1 Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, **2** Department of Health Surveillance and Management, Futian District Center for Disease Control and Prevention of Shenzhen, Guangdong, China, **3** Division of Global Health (IHCAR), Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden, **4** School of Public Health, Fudan University, Shanghai, China, **5** Institute of Public Health, Heidelberg University, Heidelberg, Germany, **6** Future Position X (FPX), Gävle, Sweden

Abstract

Background: School absenteeism is a common data source in syndromic surveillance, which allows for the detection of outbreaks at an early stage. Previous studies focused on its correlation with other data sources. In this study, we evaluated the effectiveness of control measures based on early warning signals from school absenteeism surveillance in rural Chinese schools.

Methods: A school absenteeism surveillance system was established in all 17 primary schools in 3 adjacent towns in the Chinese region of Hubei. Three outbreaks (varicella, mumps, and influenza-like illness) were detected and controlled successfully from April 1, 2012, to January 15, 2014. An impulse susceptible-exposed-infectious-recovered model was used to fit the epidemics of these three outbreaks. Moreover, it simulated the potential epidemics under interventions resulting from traditional surveillance signals. The effectiveness of the absenteeism-based control measures was evaluated by comparing the simulated datasets.

Results: The school absenteeism system generated 52 signals. Three outbreaks were verified through epidemiological investigation. Compared to traditional surveillance, the school absenteeism system generated simultaneous signals for the varicella outbreak, but 3 days in advance for the mumps outbreak and 2–4 days in advance for the influenza-like illness outbreak. The estimated excess protection rates of control measures based on early signals were 0.0%, 19.0–44.1%, and 29.0–37.0% for the three outbreaks, respectively.

Conclusions: Although not all outbreak control measures can benefit from early signals through school absenteeism surveillance, the effectiveness of early signal-based interventions is obvious. School absenteeism surveillance plays an important role in reducing outbreak spread.

Citation: Fan Y, Yang M, Jiang H, Wang Y, Yang W, et al. (2014) Estimating the Effectiveness of Early Control Measures through School Absenteeism Surveillance in Observed Outbreaks at Rural Schools in Hubei, China. PLoS ONE 9(9): e106856. doi:10.1371/journal.pone.0106856

Editor: Benjamin J. Cowling, University of Hong Kong, Hong Kong

Received: March 31, 2014; **Accepted:** August 7, 2014; **Published:** September 24, 2014

Copyright: © 2014 Fan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work has received financial support from the European Union Framework Program 7 (grant no. 241900). URL of funder's web site: <http://ec.europa.eu/research/participants/portal/desktop/en/home.html>. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: sf_nie@mails.tjmu.edu.cn (SN); gracefully@163.com (LL)

Introduction

Researchers have recently started paying attention to hysteresis in traditional surveillance systems when detecting infectious disease outbreaks. Scientists have actively explored the possibility of monitoring diseases using pre-diagnosis information since the 1990s. The theory and practice of syndromic surveillance have developed rapidly since the September 11, 2001, attacks on New York and Washington, D.C. Researchers have since found that patients' chief complaints [1], OTC medication sales [2], and

school absenteeism [3] can warn of epidemics or infectious disease outbreaks in advance.

Since use of syndromic surveillance systems has increased in developed countries [4,5], more developing countries and low-resource regions have begun to employ their own versions [6–8]. In such areas, traditional disease surveillance faces sizeable challenges: (1) residents live in crowded living environments with poor health conditions, (2) health resources are insufficient to support laboratory diagnoses, and (3) governments cannot afford sufficient education and mass grassroots training to help medical staff recognize new infectious diseases. As a result, residents in

these regions are more vulnerable to infectious disease, thereby increasing the urgent necessity for the development of a flexible, convenient, and economically viable syndromic surveillance system for developing countries and low-resource rural regions.

School absenteeism surveillance (SAS) is a common data source for syndromic surveillance. Since schools are gathering places for large groups of young children, they are typical locations for the spread of infectious diseases. School-aged children play an important role in the spread of infectious disease because they are the link between the schools and families [9]. Previous studies on SAS focused on the correlations between it and other surveillance systems [8,10–12]. This paper, however, assesses the role of SAS in three disease outbreaks (varicella, mumps, and an influenza-like illness [ILI]) during the surveillance period.

We fit an infectious dynamic model to these three real epidemics that occurred under interventions resulting from early SAS signals. Moreover, this model was used to simulate corresponding potential epidemics that could occur under the same control measures but from traditional surveillance signals defined by the “National Public Health Emergency Information Reporting and Management Specification (NPHEIRMS)” [13]. Comparing these simulated epidemic datasets generated by the dynamic model, we should be able to estimate the excess preventative effects of the control measures based on the early signals provided by SAS.

Materials and Methods

Study area

We conducted the study at 17 primary schools in Shayang County, Hubei, China, from April 1, 2012, to January 15, 2014. We applied a stratified sampling strategy to select the participant primary schools at three levels. All primary schools (2 in total) were selected at the county level in the capital of Shayang. At the township level, a random sampling was applied to obtain 25% of adjacent towns in order to reduce sampling errors as much as possible within the limited funding budget (Zengji, Shiqiao, and Hougang). In the sampled towns, all township-level schools (3 in total) were selected. At the village level, all schools (12 in total) in the sampled towns were selected (see Figure 1). The number of students in each school ranged from 88 to 2,212, for a total of 8,614 students.

Data collection

Classroom teachers collected absentee information twice daily—in the morning and at noon. If students were absent, the teachers contacted their parents in a timely manner to inquire about the cause of the absence. If the absence was due to illness, the teachers inquired about the specific disease or symptoms. Following this, the teachers summarized all absence information relevant to their classes for a data reporter before 4 PM, and then the reporter submitted the school's absence information to the SAS before 6 PM. These reports included the dates and numbers of absent students, as well as their ages, sexes, classes, addresses, reasons for absenteeism, and diseases or symptoms. In the event that timely communication with the parents was not possible, the reports lacking information concerning the reason for the absence still had to be submitted on the same day. It was possible to contact the parents later and then subsequently supplement the record with the relevant information.

Quality control

We carried out requisite knowledge and skills training for the schools' teachers. During the study period, administrators

conducted daily remote supervision and data checks through the electronic reporting system. Moreover, monthly summaries about data submission quality and feedback were provided to teachers and data reporters to ensure the maintenance of surveillance data quality.

Response to alarm signals

Once alarm signals emerged, the researchers immediately verified their authenticity using the following steps:

1. Data verification: The researchers checked for errors and duplicate data on the day of the signal, and they assessed whether the students were absent due to the same disease or symptom. In the case of errors, duplicate dates, or different absence reasons, the signal was deemed false and was discarded.
2. Primary judgment: The researchers verified the presence of a spatial correlation among the absent students on the day when the signal occurred. If a correlation was detected, the signal was deemed a suspected signal.
3. Epidemiological investigation: The suspected signal was sent to the local CDC, which then sent investigators to carry out on-site epidemiological investigations aimed at ascertaining the occurrence of an outbreak based on clinical diagnosis and the contact histories of the absent students. Once the outbreak was verified, corresponding measures were immediately taken to control disease spread.

Intervention

We formulated a standard operation procedure (SOP) of comprehensive intervention against verified outbreaks as follows:

1. Isolation: The patients were required to be quarantined either at home or in the hospital to reduce contact with others. The patients returned to school after receiving a hospital-issued post-recovery clean health certificate.
2. Disinfection: The classrooms and daily supplies were disinfected at least twice a day by wiping or spraying a solution containing 500 mg/L of chlorine with an action time of 30 minutes.
3. Ventilation: This occurred for at least 2 hours every day to keep the air fresh in all classrooms.
4. Health inspection: The teachers inspected all students for relevant symptoms in the morning and at noon. The teachers reported the manifestation of any suspected cases and immediately isolated the student for observation.
5. Provision of liquid soap: The schools provided liquid soap at the sinks on every floor to encourage students to wash their hands and to pay attention to personal hygiene.
6. Health education: We publicized disease prevention knowledge by means of broadcasts and blackboard newspapers to educate all students.

All these pre-specified control measures were taken for all three disease outbreak events in our study.

Alarm signals

School absenteeism surveillance. We set up an EARS~3Cs model as the early warning algorithm in the SAS system. This is popular among a wide variety of health departments, and it is intended for use as a cumulative sum-like method consisting of three algorithms, C1, C2, and C3, which

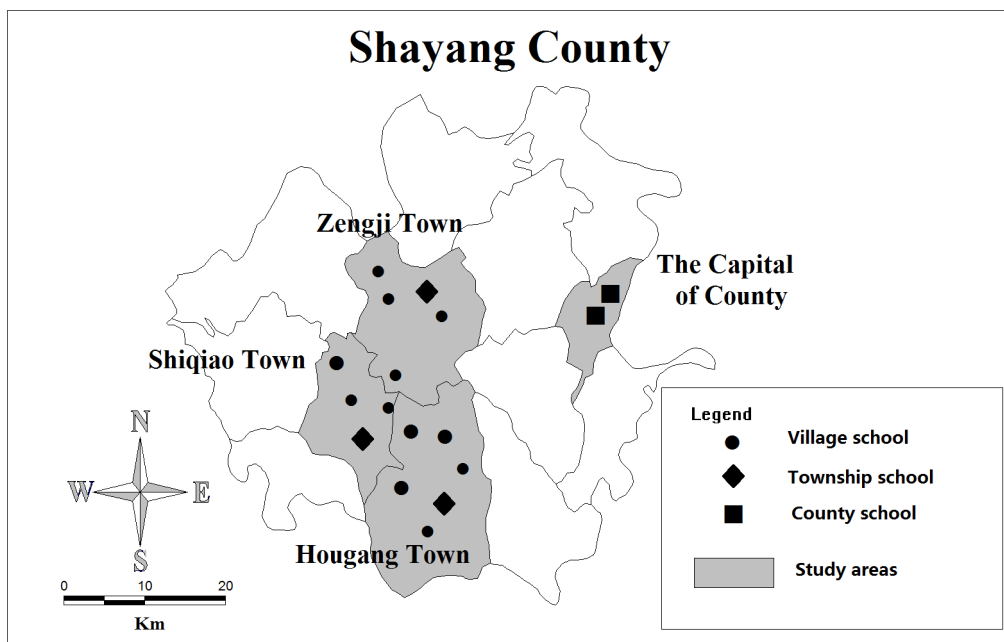


Figure 1. Geographic location of the study areas and monitoring schools.
doi:10.1371/journal.pone.0106856.g001

show increasing levels of sensitivity (C3 being the most sensitive). The C1 baseline is obtained from the previous 7 days in closest proximity to the current day (Day t-7 through Day t-1). C2 uses a 7-day baseline on Day t-9 through Day t-3. C3 is the sum of the C2 values for the past 3 days. The EARS~3C statistical values are relative numbers that reflect the fluctuation volatility of time-series data, which have taken the baseline population size into consideration. This statistic can be written as follows:

$$C1 = (x_t - \mu_1) / \sigma_1$$

$$C2 = (x_t - \mu_2) / \sigma_2$$

$$C3 = \sum_{i=t}^{t-2} \max(0, C2_i - 1)$$

where x_t is the count of cases at Day t and μ_t and σ_t are the moving sample mean and standard deviation at baseline, respectively. The EARS methodological details are described elsewhere [14].

According to the U.S. Centers for Disease Control and Prevention’s (CDC) experience with the EARS system, the threshold values of C1, C2, and C3 in this model should be set as 3, 3, and 2, respectively [14]. C1, C2, and C3 are set to automatically and simultaneously analyze the daily data for each primary school and mark the data that exceed threshold. Additionally, in order to judge the spatial associations among absentees, we set a minimum requirement of 3 absentees for an alarm signal to occur. Thus, a combined threshold was used to generate the alarm signals. If the number of school absences exceeded 3 on any given day and the marks occurred in two or more algorithm models on the same day (C1C2, C1C3, C2C3, or C1C2C3), the system sent out an alarm signal.

Traditional surveillance. China’s Ministry of Health formulated the National Public Health Emergency Information Reporting and Management Specification (NPHEIRMS) [13] in 2006. This specification defined the reporting standards for infectious disease outbreaks. All legal entities (including schools) are obligated to monitor diseases and to report possible signals in a

timely manner. Specifically, when the number of disease cases in one school exceeds the weekly reporting standard, the school has to report to the local health administrative department within two hours. Then the department carries out an on-site investigation and takes relevant outbreak control measures. According to the reporting standards of NPHEIRMS, the thresholds for the three diseases discussed in this paper are as follows: 10 varicella cases/week, 10 mumps cases/week, and 30 influenza-like illness (ILI) cases/week.

Estimation of early control measure effectiveness

The theoretical method for estimating the effectiveness of early control measures entails comparing the infection attack rate under the earlier intervention versus traditional intervention (the control group). Once early intervention is implemented, however, it is impossible to obtain the data through traditional intervention in the observed events. In order to estimate the effectiveness of early interventions, a dynamic model of infectious diseases was introduced to simulate the control group (the potential epidemic under traditional interventions) [15]. Thus, the effectiveness of early control measures can be estimated by comparing simulated epidemic datasets (generated by the dynamic model) between early and traditional interventions. The extra protective rate (EPR) can be used to evaluate the effectiveness of the early control measures. The equation for this rate is as follows:

$$EPR = (I_{e1} - I_{e2}) / I_{e1} \times 100\%$$

where I_{e1} is the estimated attack rate under the control measures based on traditional surveillance (NPHEIRMS), and I_{e2} is the attack rate under the earlier control measures based on the SAS. Both I_{e1} and I_{e2} are calculated using fitted epidemic datasets through the dynamic model.

Table 1. Parameters used in the simulation.

Parameter	varicella	Mumps	influenza
R_0 before intervention	7.0–12.0	3.8–18.2	1.6–3.0
R_0 after intervention	1.0	0.9	0.7
ω	0.08	0.09	0.5
γ	0.15	0.07	0.25
δ	0.42–0.69	0.55–0.87	0.14–0.69

δ is the possible range of positive rate of antibody in Chinese school-aged children during 2011–2013, derived from the literature.
doi:10.1371/journal.pone.0106856.t001

Impulse susceptible-exposed-infectious-recovered model

The susceptible-exposed-infectious-recovered (SEIR) model is a basic epidemic dynamic model [15]. The basic SEIR model imitates the four main health states of disease progression. Susceptible individuals (S) become exposed (E) at rate β due to contact with infected individuals. Then those exposed become infected individuals (I) at rate ω . Finally, infected individuals enter the recovery state (R) at rate γ [16,17]. The following equations model this process:

$$\left\{ \begin{array}{l} N = S(t) + E(t) + I(t) + R(t) \\ dS/dt = -\beta \times S(t) \times I(t)/N(t) \\ dE/dt = \beta \times S(t) \times I(t)/N(t) - \omega \times E(t) \\ dI/dt = \omega \times E(t) - \gamma \times I(t) \\ dR/dt = \gamma \times I(t) \end{array} \right.$$

where N is the total number of individuals. $S(t)$, $E(t)$, $I(t)$, and $R(t)$ are the numbers of individuals at time t in each of the four states. The ratios of $1/\omega$ and $1/\gamma$ are the mean incubation and infectious periods, respectively. β is the transmission rate, which reflects a disease’s diffusion intensity. This is usually measured by the reproductive number (R_0), which refers to the number of secondary cases for each primary case:

$R_0 = \beta \times (1 - \delta) \times N / \gamma$, where δ is the population’s positive antibody rate.

Unlike general surveillance systems, SAS provides discontinuous information because weekends, holidays, and vacation periods (e.g., winter and summer) fracture surveillance continuity. School class suspension largely reduces the chances for mutual student contact, and plays a notable role in hindering the transmission and spread of infectious diseases. Therefore, to improve the schools’ regular class suspensions, we added a time control variable for β into the basic SEIR model to construct an impulse SEIR model. The time control variable for β can be defined thus:

$$\beta = \begin{cases} R_0 \times \gamma / [(1 - \delta) \times N], t \neq k \\ 0, t = k \end{cases}$$

where k represents the dates schools were closed, such as weekends, holidays, or vacations.

Parameter

To fit the impulse SEIR model to the real epidemics of these three outbreaks in schools under interventions resulting from SAS signals, we defined the parameter values of R_0 before intervention, R_0 after intervention, ω , γ , and δ for varicella, mumps, and ILL.

We obtained the possible ranges of R_0 before intervention for the three diseases from those established in the existing literatures, such as 7.0–12.0 for varicella[18], 3.8–18.2 for mumps[19], and 1.6–3.0 for seasonal influenza[20].

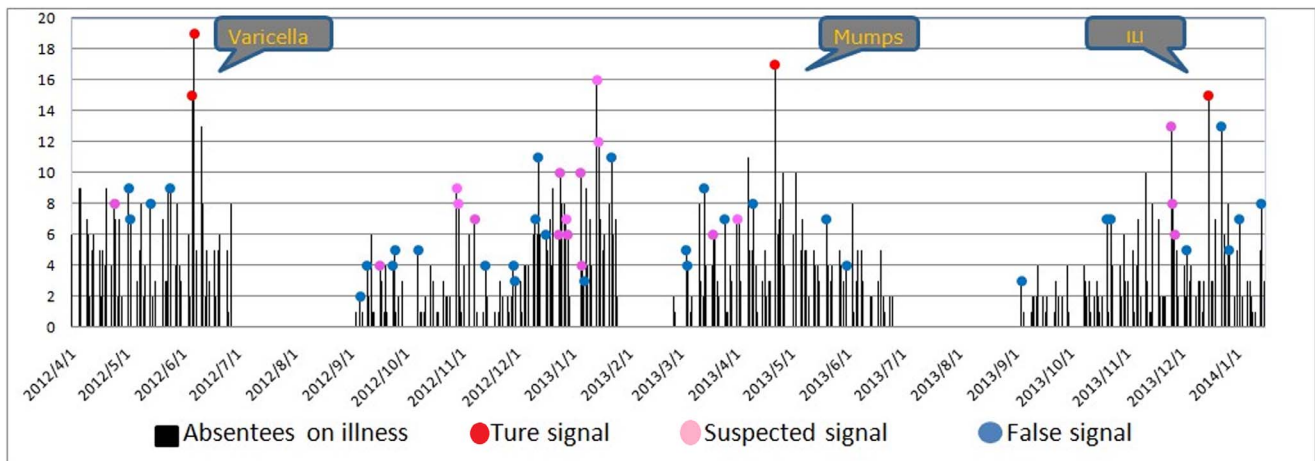


Figure 2. The three types of signals in all schools generated by school absenteeism surveillance system. Signals were generated by the EARS-3Cs.
doi:10.1371/journal.pone.0106856.g002

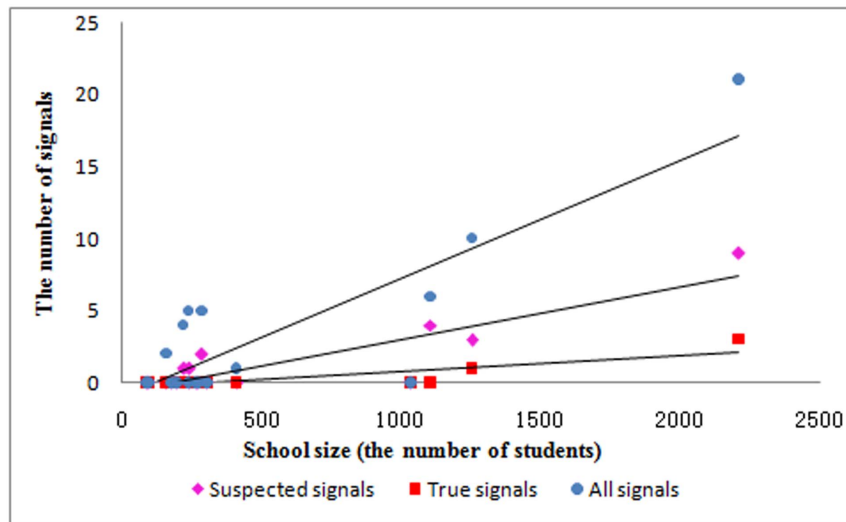


Figure 3. The correlation between school size (the number of students) and the number of signals.
doi:10.1371/journal.pone.0106856.g003

The R_0 after intervention could be estimated according to the realistic epidemic curves in each outbreak event [18]:

$$R_0 \text{ after intervention} = S_0 \times (\ln S_0 - \ln S_\infty) / (S_0 + I_0 - S_\infty)$$

where S_0 and S_∞ are the number of susceptible individuals at the start and the end of the outbreak, respectively; I_0 represents the number of cases at the beginning.

We also estimated the parameters of ω and γ , according to the average incubation and infectious periods of each disease sourced from relevant literature [21–23] (Table 1). For the study areas that were rural and resource-poor, the local CDC often lacked the ability to perform serologic tests for positive antibody rate δ among the population. Thus, we searched for relevant published data [24–30] to estimate the possible ranges of δ for the most common virus subtypes among the Chinese school-aged population from 2011 to 2013.

Ethics Statement

Written informed consent was obtained from the participants, including the children's parents. All participants and patients were de-identified, and only aggregated data was analyzed. The personal identification information did not appear in the final database. This study has been ethically approved by the Institutional Review Board of Tongji Medical College.

Results

The SAS ran for 655 days (April 1, 2012, to January 15, 2014). During this period, there were 337 (51.5%) monitoring days and 318 (48.5%) non-monitoring days (i.e., schools were closed for weekends, holidays, and vacations). A total of 1,702 (1,356 due to illness) reports were recorded, and an average of 5.1 (4.0 due to illness) students were absent from school each monitoring day. The proportion of missing data due to failure to contact parents was acceptable, accounting for only 4.8% (81/1702) of the total collected data.

By using the EARS~3Cs algorithm for absenteeism data due to illness in each school, the system detected 52 signals in 8 schools. Among these, 32 signals were excluded due to either false/duplicate records or insufficient spatial association. The staff from

the local CDC carried out epidemiologic investigations for the 20 remaining suspected signals. Consequently, 4 signals (2 varicella, 1 mumps, and 1ILI; see Figure 2 and Table 2) generated from 3 outbreak events were verified according to clinical diagnosis and contact history. In addition, Figure 3 illustrates that school size was positively correlated with the number of signals ($r = 0.86$ for all signals, $r = 0.89$ for suspected signals, $r = 0.85$ for true signals with all $P < 0.001$).

The varicella outbreak in School 17 (Figure 4a) had 41 reported cases (mean age = 8.4 years old, 18 boys) distributed among 7 classes. More than 50 percent (22) of the cases were concentrated in a single class. Signals were generated continually on June 6 and 7, 2012. The first case could be traced back to May 24, 2012. The pre-specified control measures were carried out on June 8 and lasted for 3 days.

The mumps outbreak in School 17 (Figure 4b) consisted of 28 reported cases (mean age = 9.1 years old, 15 boys) in 10 classes, with a significant proportion of (7) cases originating from a single class. The first case could be traced back to April 22, 2013. The alarm signal was noted on the same day. The pre-specified control measures were performed on April 24, 2013 and lasted for 3 days.

For the ILI outbreak in School 16 (Figure 4c), 31 reported cases (mean age = 6.9 years old, 21 boys) occurred in 8 classes, and a significant proportion of cases (13) came from a single class. The first case could be traced back to December 12, 2013, and the signal was generated on December 16, 2013. Due to the number of cases (13) that arose suddenly on this day, school administrators responded immediately, began the pre-specified control measures on the next day, and kept them going for the next 4 days.

We next fit the impulse SEIR model to the data that resulted from the school absenteeism surveillance (Figure 4). The dashes labeled "Fitted 1" are the fitted lines that were generated directly through the impulse SEIR model. Actually, SAS is unable to record the absent students on weekends due to school closure; thus the infected students who developed relevant symptoms on the weekends were recorded on the following Monday. Taking this situation into consideration, we adjusted "Fitted 1" by replacing the number of cases on weekends with "0" and adding these cases occurring on weekends with those of the following Monday ("Fitted 2"). We found that the impulse SEIR model fit the real epidemics to a large extent (Figure 4), although some systematic

Table 2. Summary of absentees, signals, and outbreaks in each monitored school.

School	N of students	N of absentees	Average absentees per day	N of absentees for illness	Average absentees for illness per day	N of signals	N of suspected signals	N of true signals	Outbreaks	Diseases
1	88	10	0.03	7	0.02	0	0	0	0	
2	99	27	0.08	17	0.05	0	0	0	0	
3	163	31	0.09	14	0.04	2	0	0	0	
4	182	35	0.10	22	0.07	0	0	0	0	
5	201	7	0.02	6	0.02	0	0	0	0	
6	223	129	0.38	86	0.26	4	1	0	0	
7	242	1	0.01	1	0.00	0	0	0	0	
8	243	55	0.16	51	0.15	5	1	0	0	
9	267	19	0.06	13	0.04	0	0	0	0	
10	277	37	0.11	20	0.06	0	0	0	0	
11	290	134	0.40	114	0.34	5	2	0	0	
12	309	5	0.01	4	0.01	0	0	0	0	
13	414	56	0.17	44	0.13	1	0	0	0	
14	1039	31	0.09	23	0.07	0	0	0	0	
15	1108	49	0.15	48	0.14	6	4	0	0	
16	1257	124	0.37	118	0.35	10	3	1	1	ILI
17	2212	952	2.82	768	2.28	21	9	3	2	varicella, mumps
Total	8614	1702	5.05	1356	4.02	52	20	4	3	

There were 337 monitoring days during the whole period. Signals were generated by the EARS-3Cs algorithm. Suspected signals were verified by data checking and spatial judgment. True signals were verified by epidemiologic investigation and clinical diagnosis.
doi:10.1371/journal.pone.0106856.t002

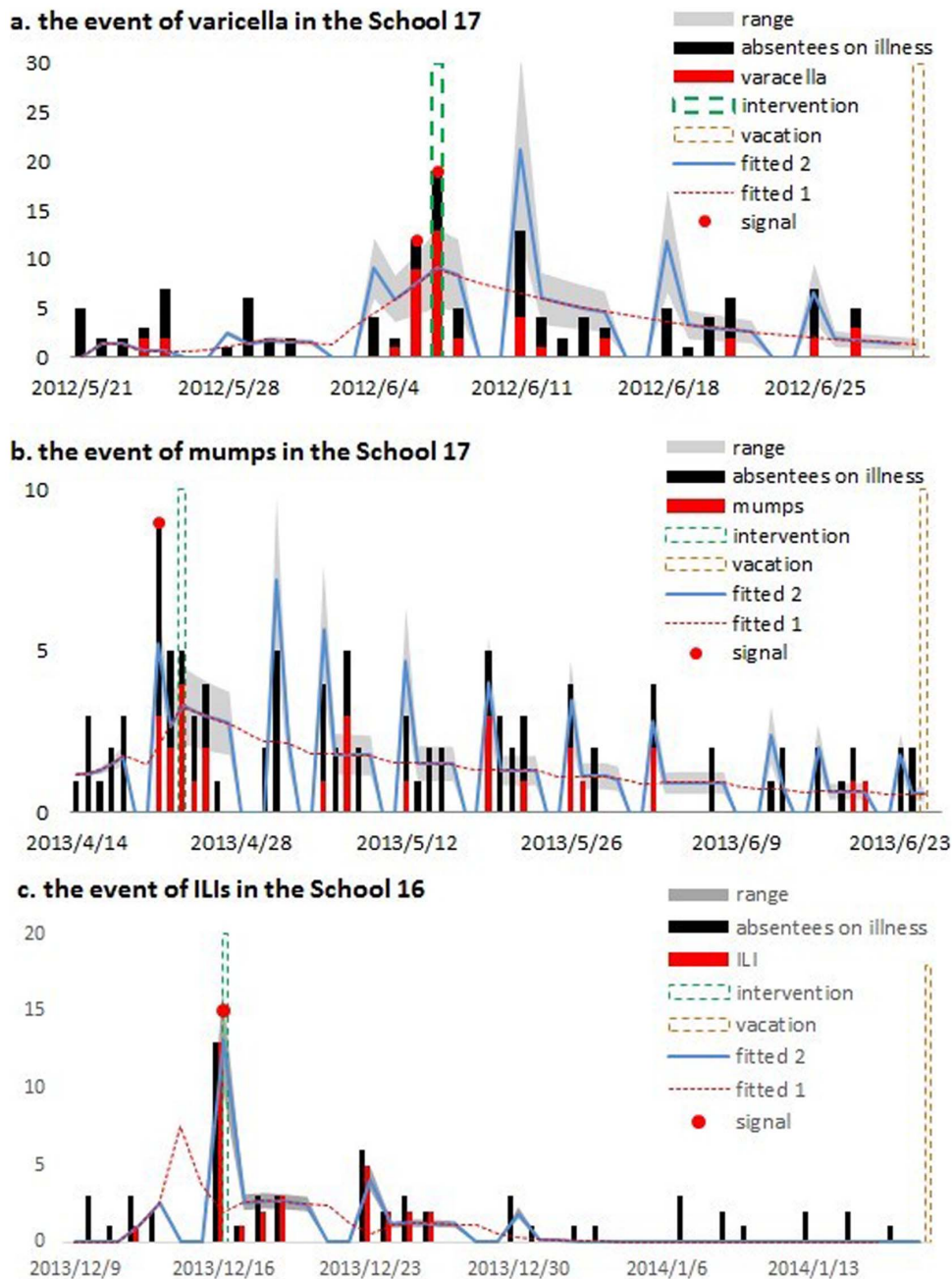


Figure 4. Model fit of epidemics in (a) varicella outbreak, (b) mumps outbreak, and (c) influenza-like illness outbreak. Vertical dash lines indicate the timing of control measures and vacations. Red dots indicate alarms by warning algorithm. 'Fitted 1' is the fitted line generated directly using impulse SEIR model; 'Fitted 2' is the fitted line adjusted by the effect of suspending reporting on weekends. doi:10.1371/journal.pone.0106856.g004

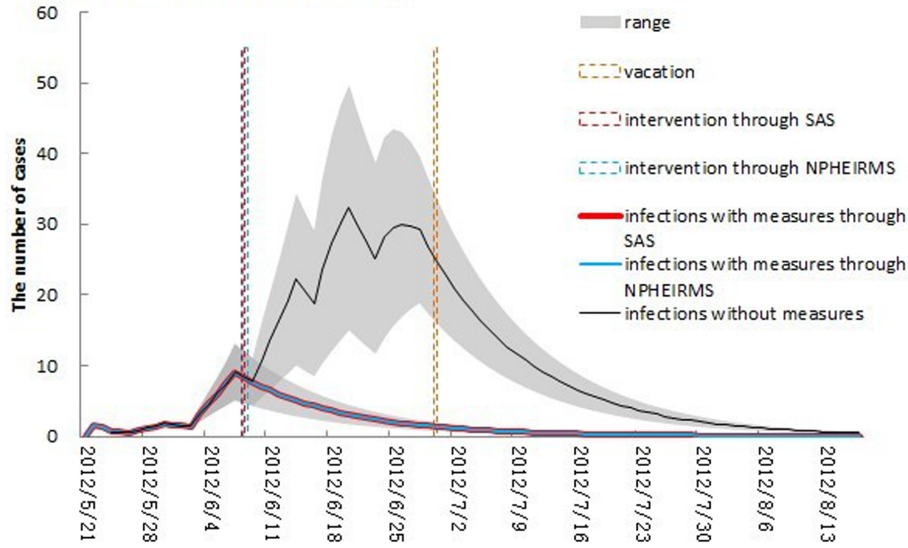
errors were likely to exist between the reported and the fitted data (e.g., students with mild symptoms may not be absent from school, particularly before the signals were generated; and biased absence reasons may be obtained in the process of enquiry for students' parents).

We simulated the potential disease epidemics without intervention (Figure 5, black lines) and with intervention based on the traditional NPHEIRMS (Figure 5, blue lines). For the varicella outbreak, the NPHEIRMS generated the signal at the same time the SAS did (the vertical dashed lines indicating the intervention timing of the SAS and the NPHEIRMS overlapped). However, for

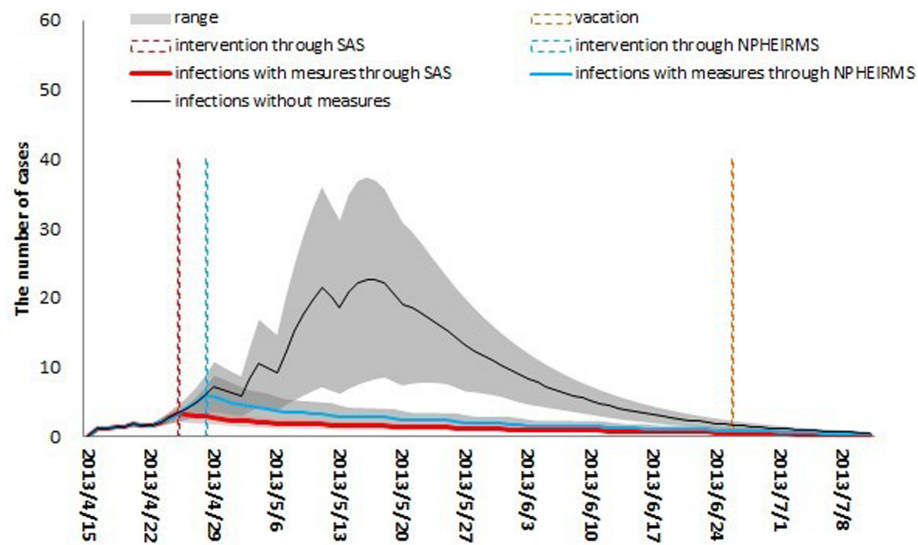
the outbreaks of mumps and ILI, the NPHEIRMS generated the signals 3 days, in average, after the SAS did. Figure 6 shows the comparisons between the timing of outbreak events based on SAS and traditional surveillance.

Using the fitted datasets of these three outbreaks under the interventions resulting from the SAS signals, we estimated the attack rates of varicella, mumps and ILI outbreaks as 7.1%–28.3%, 5.0%–41.8%, and 4.6%–14.6%, respectively. With the same alarm response time and intervention strength, the attack rates of these three outbreaks under traditional NPHEIRMS control measures would reach 7.1%–28.3%, 6.2%–62.9%, and

a. the event of varicella in the School 17



b. the event of mumps in the School 17



c. the event of Ill in the School 16

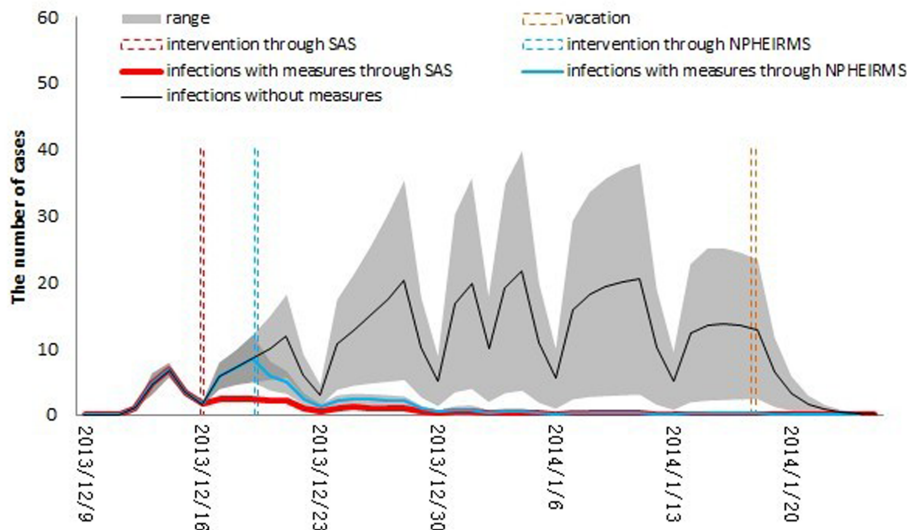


Figure 5. The simulated effects of control measures using impulse SEIR model in three outbreaks in schools. Although the surveillance for absenteeism was stopped after vacation, we still estimated the simulated infections until the epidemic faded away. SAS: school absenteeism surveillance; NPHEIRMS: National Public Health Emergency Information Reporting and Management Specification. doi:10.1371/journal.pone.0106856.g005

6.7%–22.3%, respectively (Table 3). Although the SAS failed to show obvious advantages (excess protection rate of 0.0%) compared with the NPHEIRMS for the varicella outbreak, it sent out earlier alarm signals in the mumps and ILI outbreaks. The excess protection rates indicated that the advanced interventions resulting from early signals could reduce the potential mumps and ILI cases by 15.3%–44.1% and 29.0%–37.0%, respectively.

Discussion

We analyzed three observed disease outbreaks in primary schools and compared the results with numerical simulations derived from the impulse SEIR model. Although, the SAS showed no obvious advantages in the varicella outbreak, it sent out alarms an average of three days ahead of the NPHEIRMS for the mumps and ILI outbreaks, which were effectively controlled by the early investigations and interventions of the local CDC.

In fact, the SAS’s ability to send out a warning signal ahead of the NPHEIRMS may have been linked to the early intensity of the outbreak. The current NPHEIRMS threshold values are fixed, with no fluctuations of monitoring data taken into account, and are defined according to the number of cases occurring within one week (e.g., 10 varicella cases/week, 10 mumps cases/week, and 30 ILI cases/week). If the number of cases increases rapidly during the early stage of the epidemic, the number of cumulative cases

could easily exceed the NPHEIRMS thresholds, allowing the NPHEIRMS to also send timely warning signals. However, if the number of cases increases gradually during the early stage of the epidemic, the NPHEIRMS may send a warning only when the accumulated weekly cases exceed the threshold. In this case, SAS can send out early alarm signals according to the monitored data fluctuations. Therefore, it is much more sensitive to disease outbreaks and sends out earlier warning signals much more easily.

Time is key in controlling the spread of outbreaks. An early warning signal offers the opportunity for early intervention, but whether a successful early intervention can be conducted depends on the signal’s accuracy. An early outbreak warning may be realized through syndromic surveillance based on disease-related non-specific data; but the biggest disadvantage to this is that a large number of false signals were generated. It is, therefore, essential to screen the signals sent out from any particular syndromic surveillance system. Signals can be verified by checking data quality and the spatial relationships between cases, and then undertaking epidemiological investigations.

Among the three disease outbreaks in our study, the response times differed. It is noteworthy that the measures in the ILI outbreak were taken by the school administrators themselves rather than the local CDC staff. The 13 students with influenza-like symptoms were absent on the day of the warning signal; and

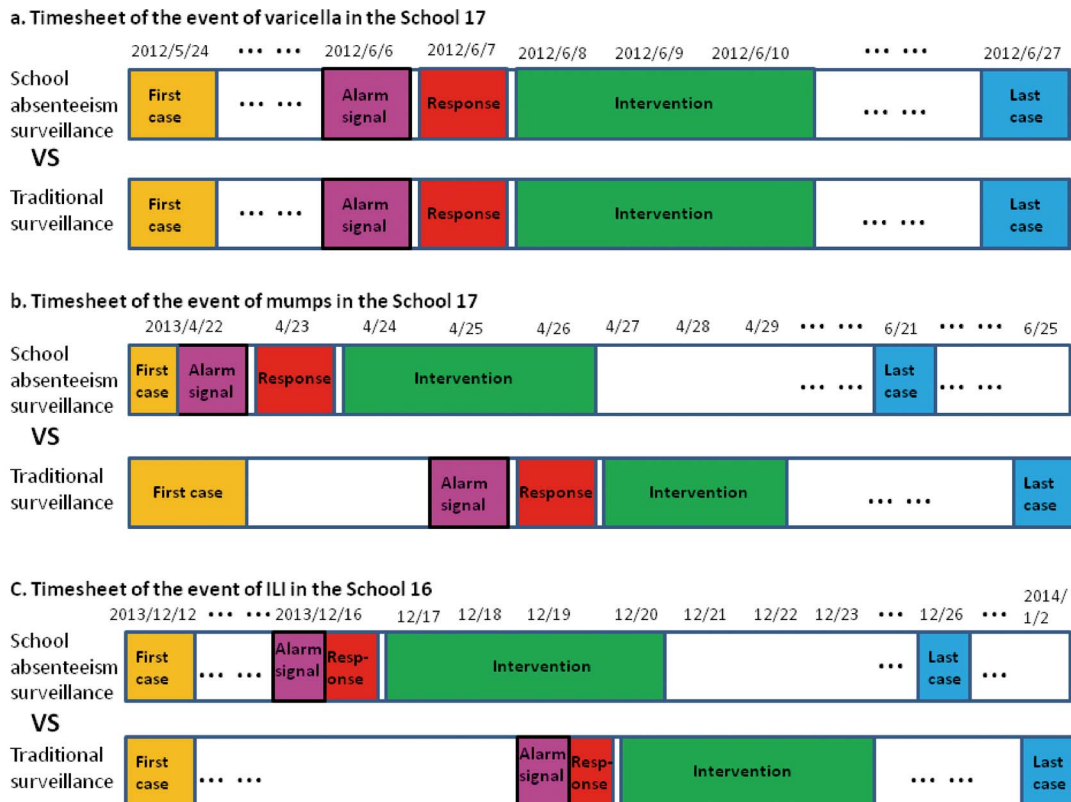


Figure 6. The comparison between the timesheets of outbreak events through school absenteeism surveillance and traditional surveillance. Time points through school absenteeism surveillance were recorded according to the actual observations, and time points through traditional surveillance were recorded according to the simulated epidemics by impulse SEIR model. doi:10.1371/journal.pone.0106856.g006

Table 3. Estimation of control measures' effectiveness through SEIR models for SAS and NPHEIRMS in three school outbreaks.

Event	R_0 before intervention	δ	N. of susceptible		SAS		NPHEIRMS		Attack rate		Time on SAS ahead of NPHEIRMS (day)	EPR of early measures through SAS (%)		
			N. of susceptible	N. of susceptible	Signal	Signal	Attack rate (%)	Attack rate (%)	N. of infections	N. of infections				
Varicella	7.0	0.42	1283	1283	2012/6/6	2012/6/6	2012/6/6	2012/6/6	7.1	7.1	91	7.1	0	0.0
		0.69	686	686	2012/6/6	2012/6/6	2012/6/6	2012/6/6	12.8	12.8	88	12.8	0	0.0
	12.0	0.42	1283	1283	2012/6/6	2012/6/6	2012/6/6	2012/6/6	16.3	16.3	209	16.3	0	0.0
Mumps		0.69	686	686	2012/6/6	2012/6/6	2012/6/6	2012/6/6	28.3	28.3	194	28.3	0	0.0
	3.8	0.55	1017	1017	2013/4/22	2013/4/22	2013/4/25	2013/4/25	5.0	6.2	63	6.2	3	19.0
		0.87	294	294	2013/4/22	2013/4/22	2013/4/25	2013/4/25	17.0	20.1	59	20.1	3	15.3
ILI	18.2	0.55	1017	1017	2013/4/22	2013/4/22	2013/4/25	2013/4/25	14.5	25.9	263	25.9	3	44.1
		0.87	294	294	2013/4/22	2013/4/22	2013/4/25	2013/4/25	41.8	62.9	185	62.9	3	33.5
	1.6	0.14	1081	1081	2013/12/16	2013/12/16	2013/12/20	2013/12/20	4.6	6.7	72	6.7	4	30.6
		0.69	390	390	2013/12/16	2013/12/16	2013/12/20	2013/12/20	12.6	17.7	69	17.7	4	29.0
	3.0	0.14	1081	1081	2013/12/16	2013/12/16	2013/12/18	2013/12/18	5.4	8.5	92	8.5	2	37.0
		0.69	390	390	2013/12/16	2013/12/16	2013/12/18	2013/12/18	14.6	22.3	87	22.3	2	34.5

SAS represents school absenteeism surveillance. NPHEIRMS is the National Public Health Emergency Information Reporting and Management Specification. EPR is the extra protective rate. doi:10.1371/journal.pone.0106856.t003

10 of these were from the same class. The school administrators attached great importance to this type of localized mass absence and took control measures themselves on the next day. School administrators with stronger syndromic surveillance awareness could speed up the response time to signals, allowing necessary additional time for early interventions. Therefore, it is helpful to carry out SAS and disease control education for both teachers and school administrators.

The key to evaluating measure effectiveness lies in how the appropriate control group is chosen. In the actual outbreak, once an intervention is performed, the surveillance system can no longer monitor the potential cases without intervention. Moreover, it is almost impossible to find another similar outbreak event to use for the control group. Therefore, the simulation of a control group carried out through an epidemic dynamic model is one of the best ways to evaluate measure effectiveness. Given the impact of weekends and holidays on school disease transmission, situations with an intervention based on traditional NPHEIRMS are simulated through the impulse SEIR model, which is in line with the characteristics of discontinuous surveillance of SAS and automatically turns the parameter of β to 0 while schools are closed, thus more realistically reflecting the school's disease transmission.

Infectious disease epidemics are closely related to population density. During the SAS surveillance period, the signals mainly occurred in large schools. The three disease outbreaks came from the two largest schools (Schools 16 and 17). Schools are the main places for social activities among school-aged children, and the increased contact among students in a comparatively closed environment greatly contributes to the spread of infectious diseases. Therefore, larger schools are more vulnerable to disease outbreaks due to more transmission agents and more frequent close contact. This suggests that SAS is more necessary in larger schools to detect possible disease outbreaks at the early stage.

There are still some problems with carrying out SAS in low-resource rural areas. First, laboratory tests for disease confirmation are not usually available in these areas. Cases can, therefore, only be confirmed through clinical diagnosis at an outbreak's earliest

stages. On the other hand, the uncertainty of virus subtypes also influences the accuracy of numerical simulation results. Second, accurate data for the seropositive rate of antibodies was not obtained, nor were sufficient funds available to test all students. Therefore, we could only use the possible ranges for seropositive antibody rates from the existing literature in order to estimate the number of susceptible individuals. Finally, although the researcher remotely supervised each school's report quality, there were still some missing reports resulting in the systematic errors between the realistic and fitted data. Therefore, increasing local teachers' awareness of syndromic surveillance in the future is also important for school absenteeism surveillance.

Although the SAS faced some challenges in the low-resource rural areas, it was able to send out early warning signals. The early control measures based on the SAS played an important role in preventing and controlling the infectious diseases, thus, the SAS can be feasibly and effectively used in rural China.

Supporting Information

Table S1 The original school absenteeism surveillance data from all 17 primary schools in the study areas. (XLS)

Acknowledgments

We thank the 17 primary schools and 3 local CDCs in the relevant study areas for providing SAS data, and for making themselves available for on-site epidemiological investigations. We also appreciate the comments of our editors and anonymous referees, which substantially improved this manuscript.

Author Contributions

Conceived and designed the experiments: YZF. Performed the experiments: HBJ YW WWY ZXZ. Analyzed the data: YZF. Contributed reagents/materials/analysis tools: WRY VKD BX HJD LP SFN. Wrote the paper: YZF LL. Provided helpful suggestions to revise the manuscript and helped respond to relevant issues from reviewers: MY.

References

1. Heffernan R, Mostashari F, Das D, Karpati A, Kulldorff M, et al. (2004) Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 10: 858–864.
2. Das D, Metzger K, Heffernan R, Balter S, Weiss D, et al. (2005) Monitoring over-the-counter medication sales for early detection of disease outbreaks—New York City. *MMWR Morb Mortal Wkly Rep* 54 Suppl: 41–46.
3. Cheng CK, Cowling BJ, Lau EH, Ho LM, Leung GM, et al. (2012) Electronic school absenteeism monitoring and influenza surveillance, Hong Kong. *Emerg Infect Dis* 18: 885–887.
4. Savage R, Chu A, Rosella LC, Crowcroft NS, Varia M, et al. (2012) Perceived usefulness of syndromic surveillance in Ontario during the H1N1 pandemic. *J Public Health (Oxf)* 34: 195–202.
5. Uscher-Pines L, Farrell CL, Cattani J, Hsieh YH, Moskal MD, et al. (2009) A survey of usage protocols of syndromic surveillance systems by state public health departments in the United States. *J Public Health Manag Pract* 15: 432–438.
6. Moore KM, Edgar BL, McGuinness D (2008) Implementation of an automated, real-time public health surveillance system linking emergency departments and health units: rationale and methodology. *CJEM* 10: 114–119.
7. Wu TS, Shih FY, Yen MY, Wu JS, Lu SW, et al. (2008) Establishing a nationwide emergency department-based syndromic surveillance system for better public health responses in Taiwan. *BMC Public Health* 8: 18.
8. Cheng CK, Channarith H, Cowling BJ (2013) Potential use of school absenteeism record for disease surveillance in developing countries, case study in rural Cambodia. *PLoS One* 8: e76859.
9. Mikolajczyk RT, Akmatov MK, Rastin S, Kretzschmar M (2008) Social contacts of school children and the transmission of respiratory-spread pathogens. *Epidemiol Infect* 136: 813–822.
10. Besculides M, Heffernan R, Mostashari F, Weiss D (2005) Evaluation of school absenteeism data for early outbreak detection, New York City. *BMC Public Health* 5: 105.
11. Suzue T, Hoshikawa Y, Nishihara S, Fujikawa A, Miyatake N, et al. (2012) The new school absentees reporting system for pandemic influenza A/H1N1 2009 infection in Japan. *PLoS One* 7: e30639.
12. Kom Mogto CA, De Serres G, Douville Fradet M, Lebel G, Toutant S, et al. (2012) School absenteeism as an adjunct surveillance indicator: experience during the second wave of the 2009 H1N1 pandemic in Quebec, Canada. *PLoS One* 7: e34084.
13. M.O.H (2006) National Public Health Emergency Information Reporting and Management Specification. Bulletin of Ministry of Health of the People's Republic of China: 44–60.
14. Fricker RD, Jr., Hegler BL, Dunfee DA (2008) Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. *Stat Med* 27: 3407–3429.
15. Andrews JR, Basu S (2011) Transmission dynamics and control of cholera in Haiti: an epidemic model. *Lancet* 377: 1248–1255.
16. Chowell G, Nishiura H, Bettencourt LM (2007) Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *J R Soc Interface* 4: 155–166.
17. Nsoesie EO, Beckman RJ, Marathe MV (2012) Sensitivity analysis of an individual-based model for simulation of influenza epidemics. *PLoS One* 7: e45414.
18. Anderson R, May R (1991) *Infectious Diseases of Human: Dynamics and Control*. Oxford: Oxford University Press.
19. Farrington CP, Whitaker HJ (2003) Estimation of effective reproduction numbers for infectious diseases using serological survey data. *Biostatistics* 4: 621–632.
20. Truscott J, Fraser C, Cauchemez S, Meechai A, Hinsley W, et al. (2012) Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza. *J R Soc Interface* 9: 304–312.
21. Heineking U, Seward JF (2006) Varicella. *Lancet* 368: 1365–1376.

22. Kutty PK, Kyaw MH, Dayan GH, Brady MT, Bocchini JA, et al. (2010) Guidance for isolation precautions for mumps in the United States: a review of the scientific basis for policy change. *Clin Infect Dis* 50: 1619–1628.
23. Cox NJ, Subbarao K (1999) Influenza. *Lancet* 354: 1277–1282.
24. Cheng R (2012) Surveillance results of antibody level of influenza among population in Futian District of Shenzhen City from 2010-2011 (in Chinese). *Occupation and Health* 28: 2942–2943,2945.
25. Ling S, Qu L (2011) Analysis on influenza surveillance and antibody level of healthy people in Longgang district of Shenzhen (in Chinese). *Progress in Microbiology and Immunology* 39: 42–44.
26. Wang T (2012) Monitoring of Rubella and Mumps Antibody Levels Among Healthy Population in Changping District, Beijing (in Chinese). *Practical Preventive Medicine* 19: 45–47.
27. Wu J (2013) Analysis of Varicella-Zoster Virus IgG Antibody Levels of Healthy Children in Jiangsu Province in 2012 (in Chinese). *Chinese Primary Health Care* 27: 87–89.
28. Yang Y (2012) Survey of antibody levels to varicella-zoster virus in healthy population in Shanghai (in Chinese). *Disease Surveillance* 27: 868–872.
29. Zhang F (2013) Monitoring and analysis of antibody levels of measles, rubella and mumps in healthy population of Zhidan County (in Chinese). *Chinese Journal of Disease Control & Prevention* 17: 1009–1010.
30. Zhang L (2012) Analysis on Immunization of Live Attenuated Varicella Vaccine in A School Student in Minhang, Shanghai (in Chinese). *Chinese Primary Health Care* 26: 81–83.