Research Report

Journal of INTERNATIONAL MEDICAL RESEARCH

Journal of International Medical Research
2018, Vol. 46(1) 98–106
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0300060517718770
journals.sagepub.com/home/imr
SAGE

Influence of infectious disease seasonality on the performance of the outbreak detection algorithm in the China Infectious Disease Automated-alert and Response System

Ruiping Wang[1,2,*], Yonggen Jiang[2,*], Xiaoqin Guo[2,*], Yiling Wu[2] and Genming Zhao[1]

Abstract

Objective: The Chinese Center for Disease Control and Prevention developed the China Infectious Disease Automated-alert and Response System (CIDARS) in 2008. The CIDARS can detect outbreak signals in a timely manner but generates many false-positive signals, especially for diseases with seasonality. We assessed the influence of seasonality on infectious disease outbreak detection performance.

Methods: Chickenpox surveillance data in Songjiang District, Shanghai were used. The optimized early alert thresholds for chickenpox were selected according to three algorithm evaluation indexes: sensitivity (Se), false alarm rate (FAR), and time to detection (TTD). Performance of selected proper thresholds was assessed by data external to the study period.

Results: The optimized early alert threshold for chickenpox during the epidemic season was the percentile P65, which demonstrated an Se of 93.33%, FAR of 0%, and TTD of 0 days. The optimized early alert threshold in the nonepidemic season was P50, demonstrating an Se of 100%, FAR of 18.94%, and TTD was 2.5 days. The performance evaluation demonstrated that the use of an optimized threshold adjusted for seasonality could reduce the FAR and shorten the TTD.

Conclusions: Selection of optimized early alert thresholds based on local infectious disease seasonality could improve the performance of the CIDARS.

[1]School of Public Health, Fudan University, Shanghai, China
[2]Songjiang Center for Disease Control and Prevention, Shanghai, China

*These authors contributed equally to this work.

Corresponding authors:
Genming Zhao and Ruiping Wang, School of Public Health, Fudan University, No. 130 Dong An Road, Xuhui District, Shanghai, China, Shanghai, 86-200032, China.
Emails: gmzhao@shmu.edu.cn; w19830901@126.com

Keywords
Moving percentile method, optimized early alert threshold, seasonality, CIDARS, infectious disease, epidemic

Date received: 21 April 2017; accepted: 12 June 2017

Background

Infectious disease remains a major public health issue in China and contributes to the high level of morbidity and mortality in the general population.[1] Detecting infectious disease outbreaks in their early stage can assist in the timely implementation of control methods.[2] In recent years, computer technology and aberration detection algorithms have achieved great developments, and they are now used to detect infectious disease outbreaks. Several national-level public health agencies have been established for infectious disease outbreak detection by automated early alert systems.[3–5] One such representative system, the China Infectious Disease Automated-alert and Response System (CIDARS), was successfully enforced in 2005 and became operational nationwide in 2008.[6] The CIDARS employs the moving percentile method (MPM), the most commonly used temporal detection method with which to detect aberration in infectious diseases. In the MRM, the reported cases in the current observation period are compared with those in a matching historical period, potentially indicating the early stages of potential outbreaks.[7]

Several studies have indicated that many determinants influence the performance of outbreak detection accuracy.[8,9] Evaluations of surveillance systems have demonstrated that the CIDARS can detect many outbreak signals in a timely manner but generates many false-positive signals, especially for infectious diseases with seasonality.[10] This indicates that the performance of infectious outbreak detection may be influenced by the epidemiologic features of the infectious disease in question. Although the CIDARS incorporates the most recent 5 years of historical data to model the influence of seasonality,[11] how seasonality influences the MPM algorithm's performance of infectious disease outbreak detection in the CIDARS remains unclear.

In this study, we evaluated the influence of seasonality on the performance of chickenpox outbreak detection. We selected the optimized early alert thresholds in the CIDARS during the epidemic and nonepidemic seasons of chickenpox and examined whether separately setting the proper thresholds according to the seasonality can improve the outbreak detection accuracy and timeliness.

Methods

Data source

The CIDARS uses the real-time individual case information housed in the Notifiable Infectious Disease Reporting Information System (NIDRIS).[6] In this study, data for chickenpox cases that occurred from 2010 to 2016 were extracted from the NIDRIS. The data were organized by day and year and saved in Excel format. Data from 2010 to 2015 were used as the baseline to select the optimized thresholds by the MPM both in the epidemic and nonepidemic seasons, while data from 2016 were used to validate the outbreak detection performance of the selected optimized thresholds.

Study design

In this study, a chickenpox outbreak was defined as five or more cases localizable to the same mass gathering, village, school, or community within 7 days.[12] Chickenpox

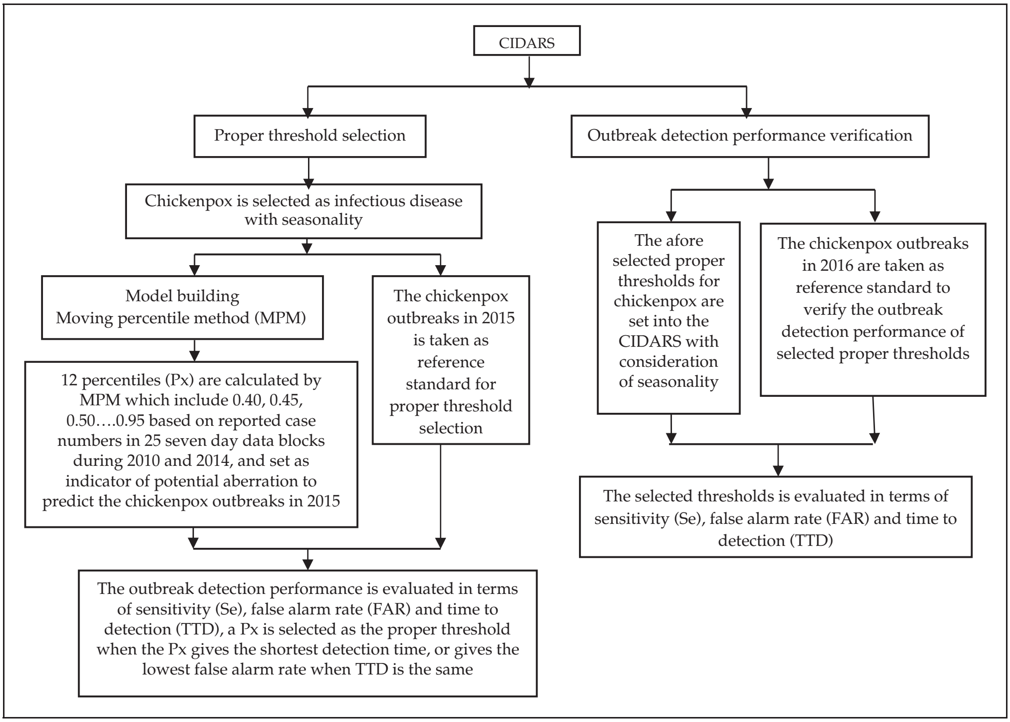

Figure 1. Flowchart of data processing and outbreak detection performance evaluation in the China Infectious Disease Automated-alert and Response System.

outbreaks reported in 2015 and 2016 in Songjiang District of Shanghai were assumed to be true outbreaks in this study. All outbreaks were investigated and verified by local staff members of the China CDC, and these outbreaks were taken as the reference standard for optimized alert threshold selection and outbreak detection performance verification.

Figure 1 shows a flowchart of the data processing in this study. The data processing involved proper threshold selection (part 1) and outbreak detection performance verification (part 2). In part 1, chickenpox was selected as the infectious disease with seasonality, 12 percentiles (Pxs) were calculated by the MPM based on the numbers of reported cases in 25 7-day data blocks during 2010 and 2014, and each Px was used to predict the actual chickenpox outbreaks that occurred in 2015. One Px was then selected as the proper threshold that indicated the shortest detection time or gave the lowest false alarm rate (FAR) when the time to detection (TTD) was the same. In part 2, selected proper thresholds with consideration of seasonality were entered into the CIDARS, and the performance of the selected thresholds for chickenpox was evaluated according to the actual outbreaks that occurred in 2016.

MPM and outbreak detection performance evaluation indexes

In the CIDARS, the MPM is applied to explore aberrations and determine the optimized threshold of a Px for a common infectious disease.[13] The MPM uses the data of the most recent 5 years as baseline data, setting the Px of the baseline data as detection parameter c. If the current day count is greater than the detection parameter's corresponding percentile (detection parameter c), then an outbreak signal is generated.[6,9] Thus, by using the MPM, aberrations in disease occurrence are detected by comparing the number of cases reported during the current observation period to that reported during a matching historical period.[13] To eliminate the weekend effect, the current observation period is defined as the most recent 7-day period, and the number of cases is the total number of reported cases in that 7-day period. The historical period is defined as the 5 years preceding the current observation year, and the matching historical period includes the same 7-day period, the two previous 7-day periods, and the two subsequent 7-day periods for each of the previous 5 years. This results in 25 7-day blocks of historical data and covers 175 days. The Pxs of the 25 historical data blocks are set as the indicators with which to detect infectious disease outbreaks. The data blocks of the current observation period and the matching historical period are moved forward dynamically day by day. If the number of cases in the current observation period is greater than the Px value of the 25 blocks of corresponding historical data, then an outbreak signal is produced.

According to previous research,[14–17] the onset of an outbreak is the onset date of the first case, and the end of an outbreak is the onset date of the last case associated with the outbreak. An outbreak is detected if a signal is generated by a Px of the MPM during the outbreak. The outbreak detection performance is evaluated in terms of the sensitivity (Se), FAR, and TTD.[16] The Se is the proportion of outbreaks detected; it is calculated as the number of detected outbreaks divided by the total number of reported outbreaks. The FAR is the proportion of early warning signals indicating false outbreaks. The TTD is the duration between the first true alarm and the onset of an outbreak; if an outbreak is flagged on the first day, then the TTD is set as 0. To calculate the timeliness of detection of all outbreaks, the total duration of an outbreak is assigned as the TTD for undetected outbreaks. To obtain the optimized early alert thresholds of MPM, 12 alternative Pxs are calculated, including P40, P45...P85, P90, and P95. A Px is selected as the optimized threshold for the MPM when that Px gives the smallest TTD or provides the lowest FAR if the TTD is the same.[17]

Optimized threshold verification

In Songjiang District of Shanghai, the incidence rate of chickenpox has seasonal characteristics; winter and early spring is the epidemic season. The seasonality of chickenpox is defined according to the epidemic curve (Figure 2), which is based on retrospective baseline data and consultation with epidemiologists. For 2015, the nonepidemic season was defined as the period from 12 February 2015 to 16 October 2015, and all other days in 2015 were defined as the epidemic season. For 2016, the period from 13 October 2016 to 31 December 2016 was defined as the epidemic season, and all other days in 2016 were defined as the nonepidemic season. The optimized early alert thresholds for chickenpox with consideration of seasonality were entered into the CIDARS. The outbreak detection performance for the optimized thresholds were verified by the Se, FAR, and TTD according to the chickenpox outbreaks reported in 2016.

Data analysis

Data analysis was performed using Excel 2013 (Microsoft, Redmond, WA, USA) and R software (version 2.14.1; R Foundation for Statistical Computing, Vienna, Austria). Excel was used to sort the data, and R was

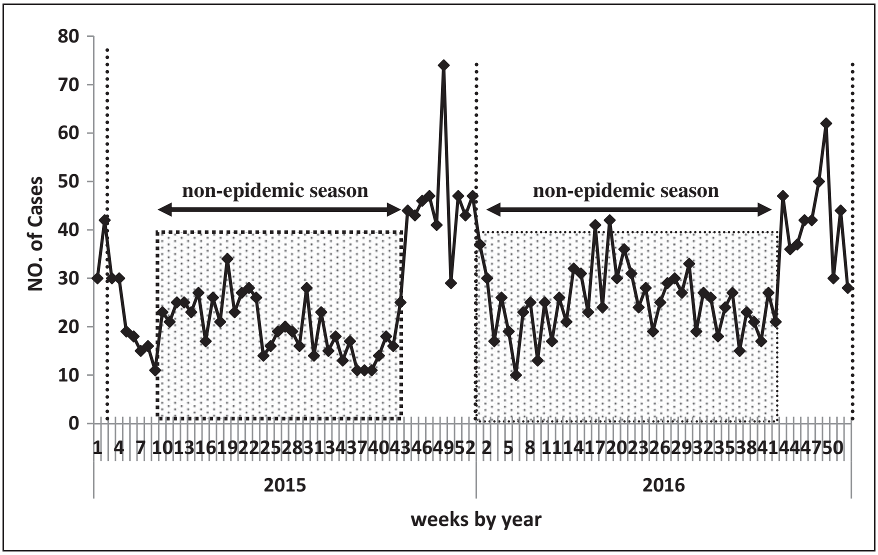

Figure 2. Chickenpox epidemic and nonepidemic season during 2015–2016 in Songjiang District, Shanghai, China.

used to calculate the evaluation indexes (Se, FAR, and TTD) and identify the proper Px threshold for chickenpox.

Results

Chickenpox incidence during 2015 and 2016

In Songjiang District of Shanghai, 1335 chickenpox cases were reported in 2015, with an incidence rate of 80.01 per 100,000. In 2016, 1484 chickenpox cases were reported with an incidence rate of 89.43 per 100,000. In both 2015 and 2016, the number of chickenpox cases reported during the epidemic season accounted more than half of that during the whole year. The incidence rate of chickenpox during the epidemic season was 51.88 per 100,000 in 2015 and 48.07 per 100,000 in 2016.

Optimized early alert threshold selection

Table 1 shows the Se, FAR, and TTD of the 12 alternative Pxs in the MPM algorithm based on the 2015 chickenpox data. Using the number of outbreaks during the whole year as the reference standard, the optimized early alert threshold was P50, demonstrating an Se of 100%, FAR of 18.98%, and TTD of 1 day. In the epidemic season, the Se was 100%, FAR was 20%, and TTD was 0 days. In the nonepidemic season, the Se was 100%, FAR was 18.94%, and TTD was 2.5 days.

When considering the influence of seasonality, the optimized early alert threshold for chickenpox during the epidemic season was P65, demonstrating an Se of 93.33%, FAR of 0%, and TTD of 0 days. The optimized early alert threshold for chickenpox in the nonepidemic season was P50, which was the same as the threshold identified without consideration of the seasonality (Table 1).

Optimized early alert threshold performance verification

Table 2 indicates that the optimized threshold of P65 in the epidemic season and P50 in the nonepidemic season for chickenpox

Table 1. Sensitivity, false alarm rate, and time to detection for 12 alternative percentiles in the moving percentile method by the whole year, epidemic season, and nonepidemic season in 2015 based on chickenpox data in Songjiang District of Shanghai, China.

	Whole year of 2015			Epidemic season in 2015			Nonepidemic season in 2015		
Alternative percentile	Se (%)	FAR (%)	TTD (days)	Se (%)	FAR (%)	TTD (days)	Se (%)	FAR (%)	TTD (days)
P40	100.00	24.82	0.5	100.00	20.00	0.0	100.00	25.00	2.5
P45	100.00	23.36	0.5	100.00	20.00	0.0	100.00	23.48	2.5
P50	***100.00***	***18.98***	***1.0***	100.00	20.00	0.0	***100.00***	***18.94***	***2.5***
P55	92.00	15.33	1.5	93.33	20.00	0.0	90.00	15.15	6.5
P60	92.00	12.41	1.5	93.33	20.00	0.0	90.00	12.12	6.5
P65	88.00	10.22	1.5	***93.33***	***0.00***	***0.0***	80.00	10.53	8.5
P70	84.00	7.30	3.0	86.67	0.00	0.0	80.00	7.52	8.5
P75	68.00	3.65	3.5	86.67	0.00	1.0	40.00	3.76	14.5
P80	44.00	0.00	15.0	60.00	0.00	3.0	20.00	0.00	21.0
P85	32.00	0.00	18.0	40.00	0.00	11.0	20.00	0.00	27.5
P90	32.00	0.00	18.0	40.00	0.00	11.0	20.00	0.00	27.5
P95	32.00	0.00	18.0	40.00	0.00	11.0	20.00	0.00	27.5

Se, sensitivity; FAR, false alarm rate; TTD, time to detection
The bold italicized text indicates the optimized early alert thresholds and corresponding evaluation indexes.

Table 2. Sensitivity, false alarm rate, and time to detection for optimal threshold by epidemic season and nonepidemic season in 2016 based on chickenpox data in Songjiang District of Shanghai, China.

Period	Optimal threshold	Signals (n)	Detected outbreaks (n)	Se (%)	FAR (%)	TTD (days)
Epidemic season in 2016	P65	23	21	100.00	8.69	2.5
Nonepidemic season in 2016	P50	69	58	100.00	15.94	0.5
Whole year of 2016	P50	107	79	100.00	26.17	3

Se, sensitivity; FAR, false alarm rate; TTD, time to detection

performed well in 2016. In the epidemic season, the Se was 100%, FAR was 8.69%, and TTD was 2.5 days. In the nonepidemic season, the optimized threshold of P50 demonstrated an Se of 100%, FAR of 15.94%, and TTD of 0.5 day (Table 2).

Discussion

The aberration detection performance of infectious disease automated surveillance systems is influenced by many determinants, and understanding how the performance is affected by these determinants can help to improve infectious disease aberration detection.[7] The results of the present study indicate that adopting an optimized early alert threshold with consideration of infectious disease seasonality can improve outbreak detection performance.

In selecting outbreak early alert thresholds, we target a minimal FAR with a high Se and short TTD.[17–20]. The results of the present study indicate that the optimized

MPM thresholds in the CIDARS differed by season for an infectious disease with seasonality, and the use of selected optimized thresholds for the corresponding seasons could significantly reduce the number of false alarm signals and detect outbreaks in fewer days during the epidemic season. The outbreak detection performance verification also demonstrated that the optimized early alert threshold could achieve a good early alert effect for infectious disease outbreak detection in both the epidemic and nonepidemic seasons.

Our findings can partly explain why infectious disease seasonality affects the outbreak detection performance of the CIDARS, particularly when considering a previous description of variation of the chickenpox incidence rates in different epidemic seasons.[17] Previous studies[20–22] have shown that for infectious diseases with seasonality, the numbers of cases and outbreaks differed greatly within different epidemic seasons; thus, the scale of outbreaks and outbreak-related case characteristics might also differ during the epidemic season.

Kuang et al.[21] reported that outbreak detection performance in automated surveillance systems is affected by many determinants. Diseases with a long incubation period have a higher Se but require a longer time for detection, and outbreaks of diseases with a short incubation time are severe but transient, leading to a lower Se. Additionally, diseases with a lower outbreak magnitude have the same Se but require more time for detection. Wang et al.[12] reported that the morbidity and mortality associated with infectious diseases and the emergency response ability of the Chinese Center for Disease Control and Prevention should also be taken into consideration during selection of the optimized early alert threshold in the CIDARS. We suggest that both the epidemic features and local characteristics of infectious diseases should also be taken into consideration. A lower threshold may be preferable if the evaluated infectious disease is associated with a tremendous threat and has reliable treatment and control measures. However, it may be wiser to select a relatively higher threshold when the infectious disease has mild effects but a high cost of investigation and control.[23]

A key strength of our study is the use of data from epidemiologically confirmed outbreaks. These confirmed outbreak data objectively reflect the real features of these outbreaks and related cases. The use of real data for optimized early alert threshold selection and performance evaluation could generate a more reliable reference standard than simulated outbreaks. Another strength of this study is that we prospectively evaluated the selection of optimized early alert thresholds for different epidemic seasons using data external to the study period.

This study also has some limitations. First, we only evaluated chickenpox as a representative infectious disease; thus, it is likely that the study results are unsuitable for other diseases with seasonality. Second, all reported chickenpox outbreaks in 2015 and 2016 were investigated and validated, but some outbreaks were inevitably missed. This may have affected the Se and FAR, especially in the nonepidemic season. Third, the selection of optimized early alert thresholds for different epidemic seasons was based on limited epidemiologically confirmed outbreaks, which may have affected the stability of the evaluation. Finally, we only evaluated the influence of the infectious disease seasonality on outbreak detection performance; however, other epidemic features including the outbreak magnitude, incubation period, and baseline counts may also influence outbreak detection performance. Future studies should incorporate improvements to fully evaluate these factors.

Conclusions

Selection of optimized early alert thresholds based on the seasonality of local infectious diseases in the CIDARS is crucial to improve the performance of outbreak detection.

Acknowledgements

We thank Professor Robert E. Fontaine (Center for Diseases Control and Prevention, US) for giving suggestions and comments for this manuscript.

Authors' contributions

RPW, YGJ, and GMZ participated in the study design. RPW conducted the study and drafted the paper. YLW and XQG revised the paper. All authors read the paper and approved the final manuscript.

Consent for publication

Not applicable.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Ethics approval and consent to participate

Not applicable.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by grants from the Key Laboratory of Public Health Safety of the Ministry of Education, Fudan University, China (GW2015-1).

References

1. Wang L, Wang Y, Jin S, et al. Emergence and control of infectious disease in China. *Lancet* 2008; 372: 1598–1605.
2. Buckeridege DL. Outbreak detection through automated surveillance: a review of the determinants of detection. *J Biomed Inform* 2007; 40: 370–379.
3. Reis BY, Kohane IS and Mandl KD. An epidemiological network model for disease outbreak detection. *PLoS Med* 2007; 4: e210.
4. Hulth A, Andrews N, Ethelberg S, et al. Practical usage of computer-supported outbreak detection in five European countries. *Euro Surveill* 2010; 15: 19658.
5. Cakici B, Hebing K, Grünewald M, et al. CASE: a framework for computer supported outbreak detection. *BMC Med Inform Decis Mak* 2010; 10: 14.
6. Yang W, Li Z, Lan Y, et al. A nationwide web-based automated system for outbreak early detection and rapid response in China. *Western Pac Surveill Response J* 2011; 2: 10–15.
7. Li Z, Lai S, Zhang H, et al. Hand, foot and mouth disease in China: evaluating an automated system for the detection of outbreaks. *Bull World Health Organ* 2014; 92: 656–663.
8. Hutwagner L, Browne T, Seeman GM, et al. Comparing aberration detection methods with simulated data. *Emerg Infect Dis* 2005; 11: 314–316.
9. Wang R, Long Y, Liu H, et al. A study of early detection for 8 communicable diseases by control graph method in Songjiang district of Shanghai, China. *Zhonghua Yu Fang Yi Xue Za Zhi* 2014; 48: 53–57. [in Chinese, English Abstract].
10. Yang WZ, Lan YJ, Li ZJ, et al. The application of national outbreak automatic detection and response system, China. *Zhonghua Liu Xing Bing Xue Za Zhi* 2010; 31: 1240–1244. [in Chinese, English Abstract].
11. Buckeridge DL, OKhmatovskaia A, Tu S, et al. Understanding detection performance in public health surveillance: modelling aberrancy detection algorithms. *J Am Med Inform Assoc* 2008; 15: 760–769.
12. Wang X, Zeng D, Seale H, et al. Comparing early outbreak detection algorithms based on their optimized parameter values. *J Biomed Inform* 2010; 43: 97–103.
13. Zhang H, Li Z, Lai S, et al. Evaluation of the performance of a dengue outbreak detection tool for China. *Plos One* 2014; 9: e106144.

14. Shumueli G and Burkom H. Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics* 2010; 52: 39–51.
15. Watkins RE, Eagleson S, Veenendaal B, et al. Applying cusum-based methods for the detection of outbreaks of Ross River virus in Western Australia. *BMC Med Inform Decis Mak* 2008; 8: 37.
16. Straetemans M, Alamann D, Eckmanns T, et al. Automatic outbreak detection algorithm versus electronic reporting system. *Emerg Infect Dis* 2008; 14: 1610–1612.
17. Li Z, Lai S, Buckeridge DL, et al. Adjusting outbreak detection algorithms for surveillance during epidemic and non-epidemic periods. *J Am Med Inform Assoc* 2012; 19: e51–e53.
18. Keramarou M and Evans MR. Completeness of infectious disease notification in the United Kingdom: a systematic review. *J Infect* 2012; 64: 555–564.
19. Ward M, Brandsema P, van Straten E, et al. Electronic reporting improves timeliness and completeness of infectious disease notification, the Netherlands, 2003. *Euro Surveill* 2005; 10: 27–30.
20. Reintjes R, Baumeister HG and Coulombier D. Infectious disease surveillance in North Rhine-Westphalia: first steps in the development of an early warning system. *Int J Hyg Environ Health* 2001; 203: 195–199.
21. Kuang J, Yang WZ, Zhou DL, et al. Epidemic feature affecting the performance of outbreak detection algorithms. *BMC Public Health* 2012; 12: 1–9.
22. Salmon M, Schumacher D, Burmann H, et al. A system for automated outbreak detection of communicable diseases in Germany. *Euro Surveill* 2016; 21. doi: 10.2807/1560-7917.ES.2016.21.13.30180.
23. Yu F, Zhang HL, Lai SJ, et al. The effectiveness of China infectious disease automated alert and response system (CIDARS) in local regions. *Zhong Hua Liu Xing Bing Xue Za Zhi* 2011; 32: 446–449.