Optimizing Use of Multistream Influenza Sentinel Surveillance Data

Eric H. Y. Lau,* Benjamin J. Cowling,* Lai-Ming Ho,* and Gabriel M. Leung*

We applied time-series methods to multivariate sentinel surveillance data recorded in Hong Kong during 1998–2007. Our study demonstrates that simultaneous monitoring of multiple streams of influenza surveillance data can improve the accuracy and timeliness of alerts compared with monitoring of aggregate data or of any single stream alone.

The use of separate data streams based on sentinel surveillance has long been an accepted approach to monitor community incidence and to enable timely detection of infectious disease outbreaks (*1*,*2*). Recently, more attention has been given to the combined analysis of multivariate sentinel data (*3*–*5*).

In this study we explored the possibility of improving the ability to more quickly detect peak periods of influenza activity in Hong Kong through simultaneous monitoring of multiple streams of sentinel surveillance data. Our findings have general implications in the choice of surveillance algorithms where multistream data are available.

The Study

The local Department of Health publishes weekly reports (*6*) from a network of 50 private-sector sentinel general practitioners (GP) and 62 public-sector sentinel general outpatient clinics (GOPC) on the proportion of patients seeking treatment for influenza-like illness (ILI), defined as fever plus cough or sore throat (*7*). In this study, we used the GP and GOPC sentinel surveillance data in 9 annual influenza seasons from 1998–1999 to 2006–2007, stratified by 4 geographic regions in Hong Kong—Hong Kong Island, Kowloon, New Territories East, and New Territories West—resulting in 8 separate data streams (Figure).

Each month a median of 1,555 specimens (interquartile range 1,140–2,740), primarily from hospitals, were sent to the Government Virus Unit of the Department of Health (*7*). We calculated the highest proportion of positive influenza isolations each season, and used these laboratory data to define the onset of each peak activity period when the proportion of positive influenza A or B isolates exceeded 30% of the maximum seasonal level (*7*).

Dynamic linear models (*8*) were used to generate alerts (online Technical Appendix, available from www.cdc.gov/EID/content/13/7/1154-Techapp.pdf). We determined that an aberration had occurred when the current observation fell outside a forecast interval generated by the model. For methods based on monitoring of single data streams only, an aberration triggers an alert. For simultaneous monitoring of all 8 data streams, we monitored separate aberrations as above and generated alerts based on the first occurrence of any aberration (M1), 2 simultaneous aberrations (M2), the first occurrence of 3 simultaneous aberrations (M3), any 2 aberrations within a 2-week period (M4), and any 3 aberrations within a 2-week period (M5). In the multistream analyses, we compared alerts produced by univariate models, which effectively assumed independence between the data streams, and multivariate models, which allowed for correlation between the data streams (online Technical Appendix).

Alerts were compared in terms of their sensitivity, specificity, and timeliness in detecting the onset of peak levels of influenza activity (*9*). We combined these metrics and estimated the area under the weighted receiver operating characteristic curve (AUWROC) as an overall measure of performance (*10*). The Table shows the highest AUWROC, for each method, from a predefined selection of parameter combinations and the sensitivity and timeliness at a fixed specificity of 95%. On the basis of aggregated data, we determined that alerts generated from the GOPC network achieved a higher AUWROC and better timeliness than those from the GP network. However, the best AUWROC from each of the data streams was produced by the GP New Territories East data, which outperformed the aggregate GP data. Conversely, for GOPC data, the performance of aggregate data was superior to that of any single data stream.

The Table also shows simultaneous monitoring results for all 8 geographic data streams from both GPs and GOPCs. For the univariate (independent) models for each data stream, methods based on simultaneous alerts perform well. The optimal methods were M2 and M3 with AUWROC of 0.89 and 0.90 and timeliness of 1.22 and 1.47 weeks, respectively, for a fixed specificity of 0.95. In general, univariate models performed better than multivariate models. Empirical correlation derived from one of the fitted multivariate models is shown in the online Technical Appendix; correlation structures under other models were similar (data not shown).

Results were insensitive to the choice of parameters (online Technical Appendix). The results also held when we varied the definition of the start of peak influenza activity between 10% and 50% of peak seasonal levels (online Technical Appendix).

*University of Hong Kong, Hong Kong Special Administrative Region, People's Republic of China

DOI: 10.3201/eid1407.080060

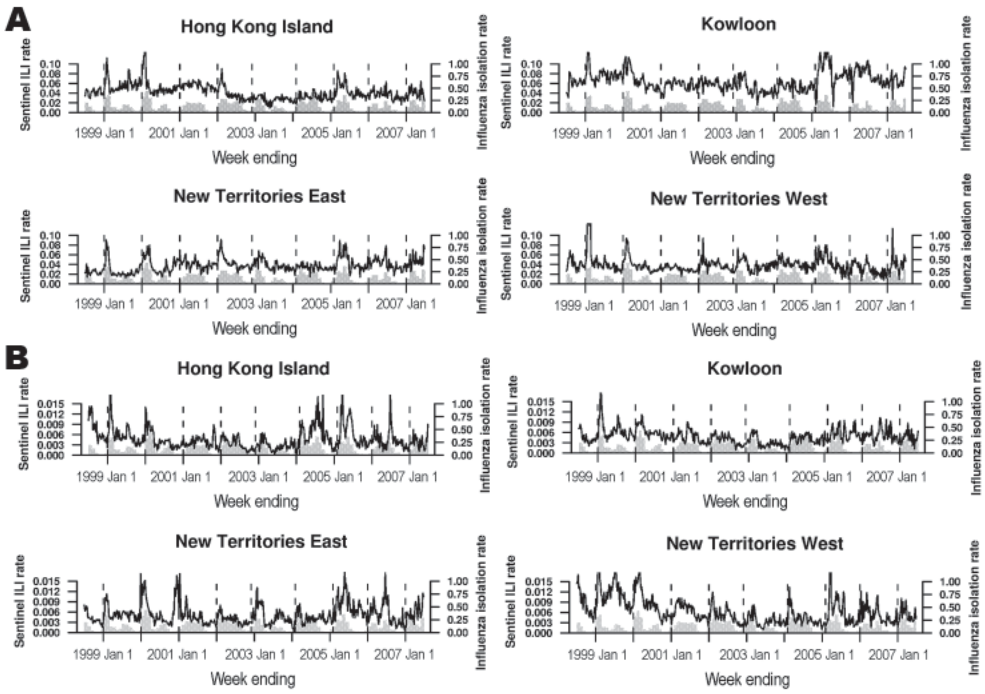

Figure. Nine annual cycles (unbroken lines) of general practitioner (A) and general outpatient clinic (B) geographic sentinel surveillance data from Hong Kong Island, Kowloon, New Territories East, and New Territories West, 1998–2007. The monthly proportions of laboratory samples testing positive for influenza isolates are overlaid as gray bars, and the beginning of each annual period of peak activity (inferred from the laboratory data) is marked with a vertical dotted line. ILI, influenza-like illness.

Conclusions

A primary objective of sentinel surveillance is to provide sensitive, specific, and timely alerts at the beginning of increased disease activity (*11*). We evaluated the performance of multistream sentinel surveillance of ILI in detecting the onset of peak influenza activity.

Splitting sentinel data into separate geographic-based streams and monitoring all 8 streams for 2 or 3 simultaneous aberrations provided substantial improvements in AUWROC and also in timeliness for a fixed specificity when compared with monitoring aggregated data or any single data stream. We also used multivariate models with various alternative correlation structures between data streams, but use of these more complex models did not appear to improve performance (Table), possibly because correlation between streams vary year to year; the multivariate model is based on constant correlations (online Technical Appendix). It is possible that other complex multivariate models may allow even greater improvement in performance; however, simultaneous monitoring of data streams may be more practical because univariate models may be applied in a spreadsheet (*7*).

Although the relative performance of GP and GOPC sentinels may not be directly generalizable to other settings with differences in infectious disease dynamics and healthcare systems, the implications for data collection are nevertheless relevant. Inclusion of data streams should be based on their value to the overall surveillance system, rather than independent performance. For example, simultaneous monitoring of data streams where some have lower specificity and others have higher specificity could still improve overall timeliness.

Specifically regarding Hong Kong, it is unclear why alerts from the private GP network have better timeliness than those from the public GOPC network. Although we note that both networks have different catchment populations, the GOPC network typically serves elderly and lower income groups (*12*), whereas influenza would be more likely to affect children at the start of the influenza season (*13*). Differences between geographic regions could be real, when disease progresses from 1 region to another (*14*); however, this circumstance is unlikely in Hong Kong, an area of only 1,000 km^2, where a high degree of mixing occurs among a population of 7 million persons. Geographic heterogeneity could also be explained by differential socioeconomics and demographics between different regions, associated differences in access to healthcare and health-seeking behavior issues, or small area variations in reporting behavior among the sentinel practices.

A potential caveat of our analysis is the small number of annual cycles of sentinel data available for study. However, until recently, few subtropical or tropical regions had begun influenza sentinel surveillance. Another limitation is the absence of a generally agreed-upon standard in defining a peak influenza season. In our analysis, the start of peak activity was defined as laboratory isolation rates exceeding 30% of the annual level; however, we found that our results were not sensitive to other reasonable thresholds. In addition, we compared methods with only a few chosen parameter combinations; sensitivity analyses showed that

Table. Performance of alerts generated by individual monitoring of aggregate data and separate data streams, and simultaneous monitoring of multiple data streams by using univariate and multivariate time series models, Hong Kong, 1998–2007*

	Univariate models			Multivariate models†		
Data	AUWROC	Sensitivity‡	Timeliness, wk‡	AUWROC	Sensitivity‡	Timeliness, wk‡
Aggregated data						
GP	0.78	1.00	2.41	–	–	–
GOPC	0.86	1.00	1.50	–	–	–
Single stream						
GP						
HK	0.75	1.00	2.36	0.73	0.87	2.64
KL	0.66	1.00	2.71	0.62	0.88	3.06
NTE	0.89	1.00	2.00	0.76	0.90	2.04
NTW	0.80	1.00	2.07	0.80	0.91	2.24
GOPC						
HK	0.79	1.00	2.21	0.71	0.89	2.42
KL	0.78	1.00	2.46	0.62	0.96	3.15
NTE	0.79	0.95	2.22	0.79	0.96	2.26
NTW	0.73	1.00	2.55	0.72	1.00	2.52
Multiple streams						
M1: First aberration	0.84	1.00	1.57	0.86	1.00	1.66
M2: 2 simultaneous aberrations	0.89	1.00	1.22	0.82	1.00	1.77
M3: 3 simultaneous aberrations	0.90	1.00	1.47	0.80	1.00	1.70
M4: Any 2 aberrations in 2 wk	0.81	1.00	2.63	0.72	1.00	2.43
M5: Any 2 aberrations in 2 wk	0.83	1.00	2.44	0.77	1.00	2.11

*AUWROC, area under the weighted receiver operating characteristic curve; GP, general practitioner; GOPC, general outpatient clinic; HK, Hong Kong Island; KL, Kowloon; NTE, New Territories East; NTW, New Territories West.
†See online Technical Appendix (available from www.cdc.gov/EID/content/14/7/1154-Techapp.pdf) for more detailed description of the multivariate model.
‡At a fixed specificity of 0.95.

the results were not sensitive to the smoothing parameter or the specification of correlations between streams. Finally, alerts generated by other more complicated combinations of aberrations might provide further enhancements. However, the value of simultaneously monitoring separate data streams (*15*) has already been demonstrated by the simple combinations chosen here.

Acknowledgments

We gratefully acknowledge the sentinel practitioners, who through their own goodwill, have been providing weekly data to the Hong Kong Department of Health for infectious disease surveillance. We also thank Irene Wong for technical assistance.

This research was in part funded by the Research Fund for the Control of Infectious Diseases of the Food and Health Bureau of the Hong Kong SAR Government (grant no. 04050102) and by the Area of Excellence Scheme of the University Grants Committee (grant no. AoE/M–12/06).

Dr Lau is a postdoctoral fellow in the School of Public Health at the University of Hong Kong. He is interested in the methodology and application of statistical and mathematical models to infectious diseases.

References

1. Besculides M, Heffernan R, Mostashari F, Weiss D. Evaluation of school absenteeism data for early outbreak detection, New York City. BMC Public Health. 2005;5:105. DOI: 10.1186/1471-2458-5-105
2. Goldenberg A, Shmueli G, Caruana RA, Fienberg SE. Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. Proc Natl Acad Sci U S A. 2002;99:5237–40. DOI: 10.1073/pnas.042117499
3. Mohtashemi M, Kleinman K, Yih WK. Multi-syndrome analysis of time series using PCA: a new concept for outbreak investigation. Stat Med. 2007;26:5203–24. DOI: 10.1002/sim.2872
4. Reis BY, Kohane IS, Mandl KD. An epidemiological network model for disease outbreak detection. PLoS Med. 2007;4:e210. DOI: 10.1371/journal.pmed.0040210
5. Fricker RD. Directionally sensitive multivariate statistical process control procedures with application to syndromic surveillance. Advances in Disease Surveillance. 2007;3:1–17. Available from http://www.isdsjournal.org/article/view/3%3B1/831
6. Hutwagner LC, Maloney EK, Bean NH, Slutsker L, Martin SM. Using laboratory-based surveillance data for prevention: an algorithm for detecting *Salmonella* outbreaks. Emerg Infect Dis. 1997;3:395–400.
7. Cowling BJ, Wong IO, Ho LM, Riley S, Leung GM. Methods for monitoring influenza surveillance data. Int J Epidemiol. 2006;35:1314–21. DOI: 10.1093/ije/dyl162
8. West M, Harrison J. Bayesian forecasting and dynamic models. 2nd ed. New York: Springer; 1997.
9. Hutwagner L, Browne T, Seeman GM, Fleischauer AT. Comparing aberration detection methods with simulated data. Emerg Infect Dis. 2005;11:314–6.
10. Kleinman KP, Abrams AM. Assessing surveillance using sensitivity, specificity and timeliness. Stat Methods Med Res. 2006;15:445–64.
11. Lawson AB, Kleinman K. Spatial and syndromic surveillance for public health. West Sussex (UK): Wiley; 2005.
12. Leung GM, Wong IO, Chan WS, Choi S, Lo SV. The ecology of health care in Hong Kong. Soc Sci Med. 2005;61:577–90. DOI: 10.1016/j.socscimed.2004.12.029
13. Glezen WP, Couch RB. Interpandemic influenza in the Houston area, 1974–76. N Engl J Med. 1978;298:587–92.

14. Alonso WJ, Viboud C, Simonsen L, Hirano EW, Daufenbach LZ, Miller MA. Seasonality of influenza in Brazil: a traveling wave from the Amazon to the subtropics. Am J Epidemiol. 2007;165:1434–42. DOI: 10.1093/aje/kwm012
15. Kleinman KP, Abrams A, Mandl K, Platt R. Simulation for assessing statistical methods of biologic terrorism surveillance. MMWR Morb Mortal Wkly Rep. 2005;54(Suppl):101–8.

Address for correspondence: Benjamin J. Cowling, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Rd, Pokfulam, Hong Kong, People's Republic of China; email: bcowling@hku.hk

Optimizing Use of Multistream Influenza Sentinel Surveillance Data

Technical Appendix

The dynamic linear model (*8*) is a form of state space model for time series data. In these models, observed data $\boldsymbol{y_t}$ (representing a vector of data at time *t*) are assumed to be related to some underlying unobserved sequential process, or system. Allowing for this underlying system where subsequent (unobserved) system values $\boldsymbol{\theta_t}$ are highly correlated, the consecutive observations $\boldsymbol{y_t}$ are assumed to be conditionally independent. With a Bayesian approach under this structure, multivariate normal distribution theory allows simple parameter estimation. Formally, the model is described by the following equations:

$$\boldsymbol{y_t} = \boldsymbol{\theta_t} + \boldsymbol{v_t} \quad \text{where } \boldsymbol{v_t} \sim N(0, \boldsymbol{V}), \text{ and}$$

$$\boldsymbol{\theta_t} = \boldsymbol{\theta_{t-1}} + \boldsymbol{w_t} \quad \text{where } \boldsymbol{w_t} \sim N(0, \boldsymbol{\Sigma}).$$

The observation error $\boldsymbol{v_t}$ and the evolution error $\boldsymbol{w_t}$ are internally and mutually independent, and the model parameters and the observation variance matrix $\boldsymbol{V}$ may be estimated from the data by using a Kalman filter (*7,8*), i.e., without requiring maximization of a likelihood function. The evolution variance matrix can be prespecified. In the case of univariate data (i.e., separate data streams), the above vectors and matrices simplify to scalars.

In our specific application where we model 8 data streams i.e., from 2 separate surveillance networks each covering 4 geographic areas, 3 types of correlation naturally arise, namely within-source correlation, within-region correlation, and residual correlation. We specified the evolution variance-covariance matrix $\boldsymbol{\Sigma}$ in the following form:

$$\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} W & A \\ A^T & W \end{pmatrix}$$

with $$\boldsymbol{W} = \sigma^2 \begin{pmatrix} 1 & \rho_W & \rho_W & \rho_W \\ \rho_W & 1 & \rho_W & \rho_W \\ \rho_W & \rho_W & 1 & \rho_W \\ \rho_W & \rho_W & \rho_W & 1 \end{pmatrix}$$

and $\boldsymbol{A} = \begin{pmatrix} \rho_I & \rho_R & \rho_R & \rho_R \\ \rho_R & \rho_I & \rho_R & \rho_R \\ \rho_R & \rho_R & \rho_I & \rho_R \\ \rho_R & \rho_R & \rho_R & \rho_I \end{pmatrix}$

Equal correlation was assumed between geographic regions for data streams from the same surveillance network, specified by the parameter ρ_W for the off-diagonal elements in $\boldsymbol{W}$. The parameter ρ_I denotes the correlation within region but under different surveillance networks, and ρ_R denotes residual correlation between data streams in different regions and sources. Several sets of correlation parameters were prespecified to assume different levels of dependence between data streams, assuming that $\rho_W > \rho_I > \rho_R$. We applied the multivariate model using sets of (ρ_W, ρ_I, ρ_R) = (0, 0, 0), (0.3, 0.1, 0), (0.5, 0.2, 0.1), (0.8, 0.2, 0.1) and (0.8, 0.4, 0.3), denoted by $\boldsymbol{\Sigma}_i$, i =1, 2, …, 5, respectively, to represent a spectrum of assumptions in the correlation structure between data streams (from no or low inter and intradependence within source or region to high dependence). The overall evolution variance σ^2 is also varied, in combination with different dependence structure, as 0.025, 0.050, 0.075 and 0.100 to represent high- to-low smoothness in the assumed change in influenza prevalence.

Having fitted the model for data at times 0,…, *t-1*, a one-step-ahead forecast interval may be estimated for time *t*. Varying the coverage of the forecast interval provides a range of values for the 3 metrics of sensitivity, specificity, and timeliness. Increasing the width of the forecast interval will result in fewer aberrations (an aberration occurs when an observed value exceeds the forecast interval) and, in general, will increase the specificity but decrease the sensitivity and timeliness.

For simultaneous monitoring of separate data streams based on univariate models, a dynamic linear model was fitted separately to each data stream and individual aberrations were combined to generate overall alerts. This approach is very similar (but not quite identical) to assuming $\rho_W = \rho_I = \rho_R = 0$ in a multivariate model.

Estimating the AUWROC

The area under the weighted receiver operating characteristic curve (AUWROC) is a combined metric of sensitivity, specificity, and timeliness. First, the time saved (TS) is calculated as 1 minus the ratio of timeliness to a predefined maximal delay of alerts. In this scenario of influenza surveillance, considering the short generation time of the disease, an

early warning should be issued ideally within ≈4 weeks of the start of the peak season; therefore, we used 4 weeks as the maximal delay of alerts. By setting different thresholds for each method to obtain sets of sensitivity, specificity, and timeliness, a weighted ROC curve was then plotted by TS × sensitivity against the false positive rate (1 – specificity); the area under the curve is the AUWROC.

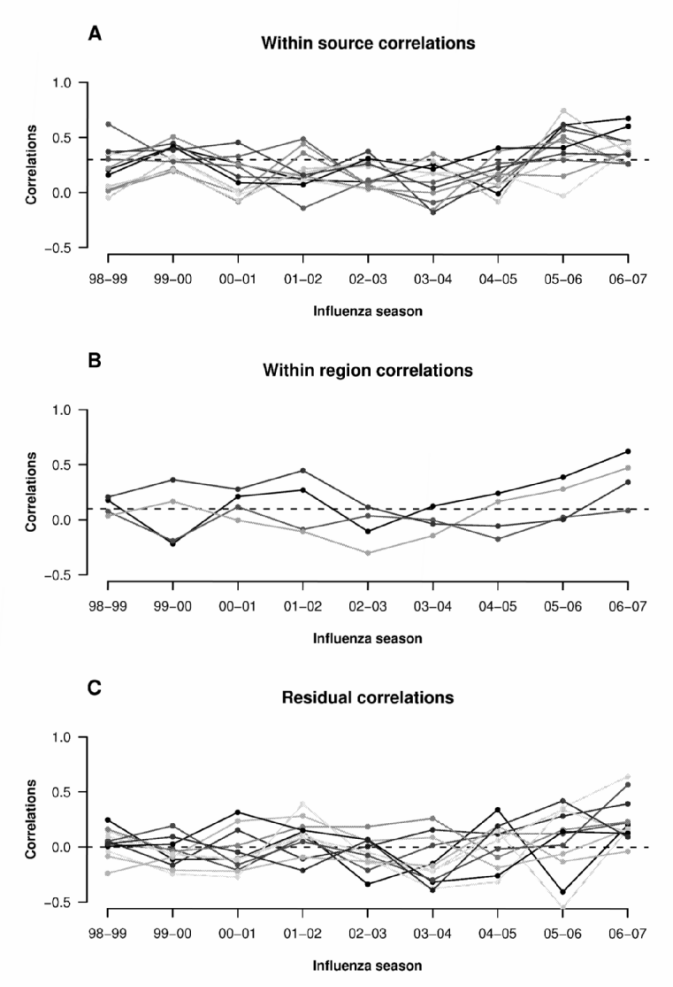

Figure. Empirical evolution correlations obtained from the dynamic linear model with parameters σ = 0.75, (ρ_W, ρ_I, ρ_R) = (0.3, 0.1, 0) over 9 influenza seasons including within-source correlations (A) within-region correlations (B), and residual correlations (C) with the horizontal lines showing the respective assumed correlations in the model.

Table 1. Sensitivity analyses of the performance of alerts generated by simultaneous monitoring of multiple data streams by using univariate and multivariate time series models*

			AUWROC				
	Parameters	Univariate	Multivariate models (ρ_W, ρ_I, ρ_R)				
Method	σ^2	models	(0, 0, 0)	(0.3, 0.1, 0)	(0.5, 0.2, 0.1)	(0.8, 0.2, 0.1)	(0.8, 0.4, 0.3)
M1: first aberration	0.025	0.84	0.84	0.82	0.81	0.81	0.81
	0.050	0.83	0.84	0.83	0.81	0.81	0.81
	0.075	**0.84**	0.85	0.83	0.81	0.82	0.81
	0.100	0.84	**0.86**	0.83	0.82	0.83	0.83
M2: 2 simultaneous aberrations	0.025	0.87	0.75	0.76	0.76	0.79	0.80
	0.050	0.88	0.78	0.80	0.79	0.80	0.80
	0.075	**0.89**	0.79	0.80	0.81	0.81	0.80
	0.100	**0.89**	**0.82**	0.80	0.81	0.81	0.81
M3: 3 simultaneous aberrations	0.025	0.89	0.74	0.75	0.74	0.74	0.77
	0.050	0.89	0.76	0.75	0.74	0.75	0.77
	0.075	0.89	0.75	0.78	0.76	0.77	0.77
	0.100	**0.90**	0.77	**0.80**	0.77	0.77	0.80
M4: any 2 aberrations in 2 wk	0.025	0.79	0.70	0.67	0.68	0.71	0.71
	0.050	0.78	0.68	0.70	0.70	0.70	0.71
	0.075	0.79	0.71	0.71	0.69	**0.72**	0.71
	0.100	**0.81**	0.71	0.67	0.70	0.70	0.71
M5: any 3 aberrations in 2 wk	0.025	**0.83**	0.67	0.70	0.69	**0.77**	0.73
	0.050	0.83	0.67	0.73	0.70	0.71	0.70
	0.075	0.82	0.69	0.72	0.71	0.71	0.70
	0.100	0.83	0.72	0.69	0.74	0.70	0.70

*AUWROC, area under weighted receiver operating characteristic curve. Optimal parameter combinations for each method under univariate and multivariate models are shown in **boldface**.

Table 2. AUWROC of different methods with the start of influenza season defined as influenza isolation rates exceeding 10%, 30%, or 50% of the seasonal peak level*

	AUWROC/threshold					
	10%		30%		50%	
Data	Univariate	Multivariate	Univariate	Multivariate	Univariate	Multivariate
Aggregated data						
GP	0.79	–	0.78	–	0.84	–
GOPC	0.76	–	0.86	–	0.85	–
Single stream						
GP						
HK	0.60	0.62	0.75	0.73	0.77	0.81
KL	0.75	0.62	0.66	0.62	0.76	0.75
NTE	0.78	0.69	0.89	0.76	0.91	0.83
NTW	0.82	0.77	0.80	0.80	0.87	0.85
GOPC	0.61	0.64	0.79	0.71	0.88	0.80
HK						
KL	0.68	0.61	0.78	0.62	0.85	0.76
NTE	0.72	0.70	0.79	0.79	0.79	0.74
NTW	0.80	0.70	0.73	0.72	0.83	0.80
Multiple streams						
M1: first aberration	0.83	0.84	0.84	0.86	0.89	0.88
M2: 2 simultaneous aberrations	0.77	0.77	0.89	0.82	0.90	0.88
M3: 3 simultaneous aberrations	0.77	0.72	0.90	0.80	0.92	0.90
M4: Any 2 aberrations in 2 wks	0.85	0.77	0.81	0.72	0.84	0.82
M5: Any 3 aberrations in 2 wks	0.83	0.75	0.83	0.77	0.87	0.86

*AUWROC, area under weighted receiver operating characteristic curve; GP, general practitioner; –, not applicable; GOPC, general outpatient clinic; HK, Hong Kong Island; KL, Kowloon; NTE, New Territories East; NTW, New Territories West.