衞生防護中心
Centre for Health Protection

Working Group on Influenza-like Illness Statistical Detection of Aberrations (WGILI)
Submitted to
The Scientific Committee on Advanced Data Analysis and Disease Modelling (SCADADM)

Working Group Report on A Comparison of Methods for Early Detection of the Influenza Peak Season in Hong Kong

Summary

What is already known? Different methods exist for automated monitoring of surveillance data, and no method is universally superior.
What does this report add? This report finds that time series methods are superior to regression and CUSUM methods for early detection of the influenza peak season in Hong Kong using sentinel surveillance data, in terms of sensitivity, specificity and timeliness.
What is the public health/policy relevance? Using time series methods on the sentinel surveillance data will allow a highly reliable alert to be made within 1-2 weeks of the start of the peak season.

Introduction

An important function of public health surveillance is to detect 'aberrations' (which may be unusual clusters or outbreaks of disease, or onsets of epidemics) and initiate timely interventions. Recent developments in computer-assisted outbreak detection offer a range of approaches to infectious disease monitoring.

2. In Hong Kong, sentinel general practitioners report weekly numbers of consultations with patients with influenza-like-illness (defined as fever plus cough or sore throat), as well as the total number

of consultations. These data may be used to monitor for the start of the influenza peak season, which generally starts in January, February or March (1).

3. It would be useful to have a valid and reliable way to predict the onset of this peak season, so that advanced warning can be given to health care providers with the objective of enhancing case detection and diagnosis for health care planning purposes, and to implement necessary precautionary measures in vulnerable populations such as the elderly (2).

4. In this report, we first describe the approaches which are available for automated monitoring of sentinel surveillance data. We then present and discuss the results of a comparison of methods.

What data are available?

5. Since 1998, weekly influenza surveillance data have been provided by two groups of sentinels, namely general practitioners (GPs) and general outpatient clinics (GOPCs). These sentinels report the weekly number of consultations with patients complaining of symptoms of influenza-like-illness (defined as fever plus cough or sore throat), as well as the total number of consultations in the preceding week. These data can be used to calculate the weekly rate of ILI consultations per 1000 total consultations.

6. We investigated both sources of sentinel surveillance data, and found that the GP data followed a smoother pattern outside the peak season, with less random variability around a mean of approximately 40 ILI consultations per 1000 total consultations. Furthermore, a (second) summer peak in influenza consultations was clearly visible in the annual cycles of GOPC data, but not in the GP data.

7. Since the GP data are smoother, and show a longer 'baseline' period before the start of the peak season, it seems sensible to focus on these data for predicting the start of the annual peak season.

8. To decide when the influenza peak season actually started, we looked at the patterns in the sentinel data, and also at virologic data provided by the Government Virus Unit of the Department of Health. For 2001-2004, weekly numbers of influenza A and B were available, and these confirmed our choice of cut-offs for the start of the peak season.

What methods are available?

9. There are a number of different approaches for the detection of aberrations in surveillance data. A common group of methods is based on a regression model originally proposed by Serfling (3). In this approach, a sinusoidal curve is fitted using cleaned historical data, and the alarm is raised if

the observed number of events falls outside a calculated boundary for two consecutive weeks. Related regression methods have been used in France, the UK CDSC (4), and ISIS in the Netherlands.

10. An approach that is growing in popularity for automated surveillance is based on CUSUMs (5), which were originally developed for industrial quality control. These methods may incorporate historical information (6) or be based on short-term data (7).

11. Surveillance data naturally form a time series, and a variety of methods have been developed for the analysis of this form of data, for example ARIMA models from the Box Jenkins family (8), and dynamic linear models (9). Reis et al. (10) propose a hybrid method which uses CUSUMs on top of an ARIMA model.

Is there any relevant literature?

12. There have been few quantitative comparisons of the different approaches, for specific applications. A recent study by the CDC (7) suggests that non-historical (i.e. short-term) regression and CUSUM approaches are superior to regression and CUSUM models with historical data, across a wide range of syndromes. However, this study does not include other approaches, and does not discuss which methods are most appropriate for early detection of the onset of an annual influenza epidemic. No other studies have been found which quantitatively compare the different approaches discussed here. Without such a quantitative comparison, it is not possible to say which approach will perform better for early detection of the annual peak in influenza infections in Hong Kong.

Which approaches have we compared?

13. Without long-term historical data, it makes sense to focus on short-term methods. Thus, we compare three short-term approaches, namely a time series method (9), a regression method (4) and a CUSUM method (6). For each approach, we choose forty different parameter combinations, aiming to give a broad range of characteristics in terms of sensitivity, specificity and timeliness as defined below.

How have we compared the approaches?

14. Six cycles of sentinel surveillance data are available, shown in Figure 1 below. It is also possible to characterise these sentinel data in terms of ‘signal+noise’, where the ‘signal’ is a peak that is added to represent the peak season, and the ‘noise’ may contain a linear trend in addition to random variation.

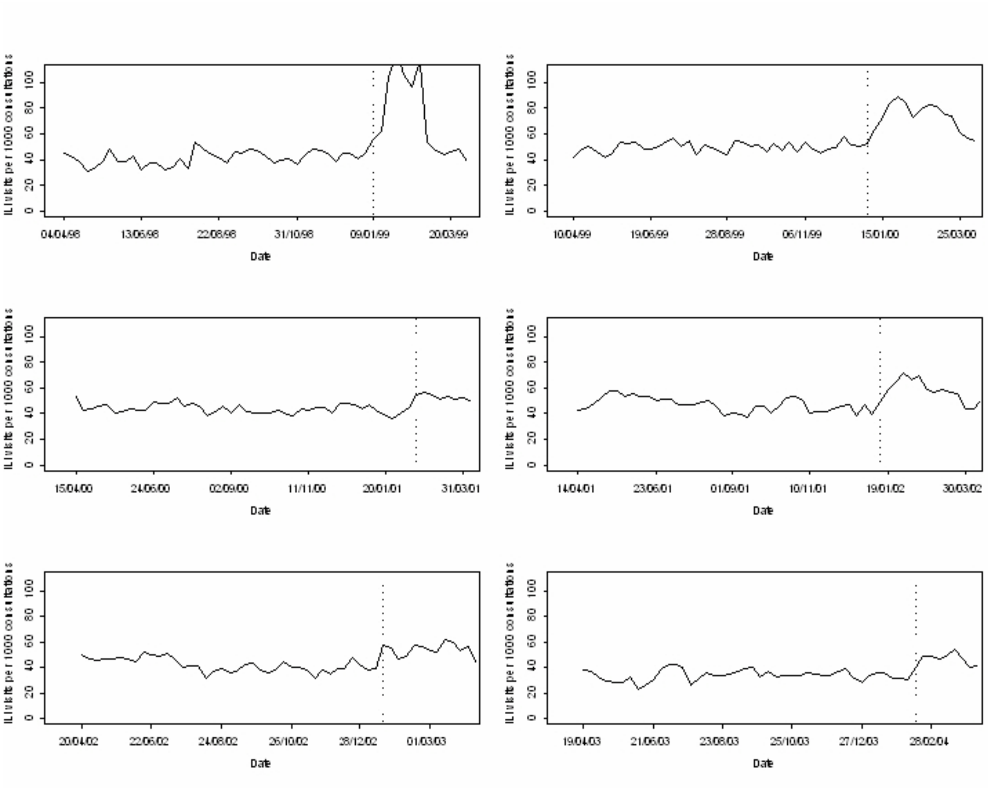

Figure 1: Six cycles of sentinel surveillance data from general practitioners in Hong Kong. The dotted line marks the start of the annual peak season.

15. This characterisation was used to generate 1000 artificial cycles of data.

16. Following Hutwagner et al. (7), we can calculate the sensitivity (the proportion of peak seasons detected), the specificity (1-*r*/*m* where *r* is the number of false positive alarms and *m* is the total number of non-peak season weeks), and the timeliness (the number of weeks before an alarm is generated after the start of the peak-season). These quantities can be calculated for each method, and compared. The ideal method will have high sensitivity, high specificity, and low timeliness.

17. In Hong Kong, the sentinel data are generated from a Monday to a Friday and reported over the weekend for collation and analysis by CHP on the following Monday. For the time to detection calculation, week 0 is defined as the Monday following the first week of the peak season, and week 1 is the Monday after that, and so on. For a particular method in a particular annual cycle, the time to detection is simply the week that the peak season was detected. Thus the mean time to detection can take any value greater than or equal to zero, and a mean time to detection of zero would indicate that the peak season is always detected as quickly as possible, given the reporting delay.

18. With forty different parameter combinations for each of three approaches, a total of 120 'methods' are compared on the six empirical cycles and then on the 1000 artificial cycles.

What were the results?

19. Table 1 below shows a summary of the characteristics of the most optimal parameter combination for each approach. The full results showed that the time series approach had better performance than the regression approach. The CUSUM approach had the worst performance of the three methods.

20. The timeliness (mean time to detection) for the time series method was 0.5 and 0.4 on empirical and simulated data respectively. When looking at the full results, we found that the time series method nearly always detected the peak season in week 0 or in week 1, and rarely in weeks 2 or later. Allowing for the reporting delay of between 3 and 7 days (depending on which day of the week prior to week 0 the peak season actually started), we can say that the time series method nearly always detected the peak season within 1-2 weeks of it starting.

Table 1: Results of empirical and simulation study.

Approach	Empirical study			Simulation study		
	Sensitivity	Specificity	Timeliness (weeks)	Sensitivity	Specificity	Timeliness (weeks)
Time series	1.000	0.989	0.500	0.996	0.968	0.405
Regression	1.000	0.968	0.500	0.964	0.976	0.519
CUSUM	1.000	0.935	0.667	0.964	0.975	0.552

Discussion

21. The results of the empirical analysis and the simulation study both show that the time series approach is superior to the other two methods under consideration. It is expected that other methods which can dynamically track the level of the data, such as Box Jenkins models, will also perform well on these sentinel surveillance data. However, the results also show that the regression approach can work well, for selected optimal parameter choices.

22. A limitation of this study is that it is based on only six annual cycles of empirical data from the sentinel practitioners. However, the results of the simulation study seem to validate the empirical results. A limitation of the data is the reporting delay, which implies a delay of at least 3 days in the discovery of the onset of the peak season. In future, it might be possible to shorten this reporting delay, for example by encouraging the sentinels to send data on the Friday evening at the end of each week, or by switching to daily reporting (perhaps using the internet).

23. In Hong Kong, other sources of influenza surveillance data have recently started to provide reports, and future research could investigate a decision rule based on many streams of surveillance data, for alerting the start of the peak season. Another interesting question for future research relates to the best definition of influenza-like-illness. The Hong Kong sentinels currently use the definition ‘fever plus cough or sore throat’, but an alternative definition might produce more informative data for sentinel purposes.

Implementation

24. A spreadsheet containing the formulas for the time series approach has been provided to the CHP.

Centre for Health Protection
March 2005

References

1. Yap, F.H.Y., P.L. Ho, K.F. Lam, P.K.S. Chan, Y.H. Cheng, and J.S.M. Peiris, Excess hospital admissions for pneumonia, chronic obstructive pulmonary disease, and heart failure during influenza seasons in Hong Kong. Journal of Medical Virology, 2004. **73**: p. 617-23.
2. Quigley, C., W. Sopwith, and M. Ashton, How to deal with influenza: Worthwhile surveillance system is in action. BMJ, 2004. **329**(7476): p.1239.
3. Serfling, R.E., Methods for current statistical analysis of excess pneumonia-influenza deaths. Public Health Reports, 1963. **78**: p. 494-506.
4. Farrington, C.P., N.J. Andrews, A.D. Beale, and M.A. Catchpole, A statistical algorithm for the early detection of outbreaks of infectious disease. Journal of the Royal Statistical Society, Series A, 1996. **159**(3): p. 547-63.
5. O'Brien, S.J. and P. Christie, Do CuSums have a role in routine communicable disease surveillance? Public Health, 1997. **111**(4): p. 255-8.
6. Hutwagner, L.C., E.K. Maloney, N.H. Bean, L. Slutsker, and S.M. Martin, Using laboratory-based surveillance data for prevention: an algorithm for detecting Salmonella outbreaks. Emerging Infectious Diseases, 1997. **3**(3): p. 395-400.
7. Hutwagner, L.C., T. Browne, G.M. Seeman, and A.T. Fleischauer, Comparison of aberration detection methods using simulated data. Emerging Infectious Diseases, 2005. **11**(2): p. 314-6.
8. Chatfield, C., Time-series forecasting. 2001, Boca Raton: Chapman & Hall.
9. West, M. and J. Harrison, Bayesian forecasting and dynamic models. 2nd ed. 1997, New York: Springer.
10. Reis, B.Y., M. Pagano, and K.D. Mandl, Using temporal context to improve biosurveillance. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(4): p. 1961-5.