Reviewer's report

Title: Establishing a Nationwide Emergency Department-based Syndromic Surveillance System for Better Public Health Responses in Taiwan

Version: 4  Date: 18 June 2007

Reviewer: Richard Heffernan

Reviewer's report:

General

This paper describes Asia’s first nationwide emergency department syndromic surveillance system. The authors document various operational aspects of the system and present graphs of daily and weekly time series for various syndrome counts, aggregated at the national level. The paper is interesting as an example of the international adoption of syndromic surveillance, and it is important that systems outside the united states and europe be published literature. Nevertheless clarification is needed in a number of places and the manuscript would benefit from more critical analysis of the public health value of this system.

----------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

In several places the authors express optimism about the capabilities of syndromic surveillance without providing supporting evidence. Examples include:

• p7 last sentence of results: “These algorithms will have many practical uses in monitoring daily changes of infectious diseases in Taiwan.”
• p8 “Syndromic surveillance can be very helpful in the early detection of unusual phenomena or fast-spreading infectious diseases”
• p10 “ED-SSS can play an important role in detecting an outbreak [of avian influenza] even if the cluster size is small.”
• p10 the ED-SSS is “crucially important” for addressing “dengue, malaria, AIDS and SARS”

The field of syndromic surveillance has matured to that point that readers expect more critical examination of the evidence that systems are valuable. The authors could provide additional specific examples where their system has proven useful, should consider extending the study period beyond the first year of operation to include more recent successes, and should include discussion of the limitations of the system.

The authors report “striking” day-of-week and holiday effects on total ED visit volume, evident in Figure 4. These are common features of health utilization data and one of the few sources of variability that can readily be controlled for. Rather than highlighting this as a “unique” characteristic of their system they should apply some kind of adjustment technique (e.g. simply track the proportion of syndrome visits to total visits, or use linear regression). Removing this artifact will help reveal illness trends.

On p6 the authors provide information re. data completeness, however some important questions are not addressed. In particular they should indicate how complete the data are at the time of daily, prospective analysis. What proportion of hospitals report on time? What proportion of ED visits are reported on time? What proportion of key fields such as ICD9 codes, chief complaint, age and zip code fields are complete at the time of daily analysis? Some of the information they do provide is ambiguous. For example, they state that only 26% of hospitals sent data every day for the first four months of operation. This could be consistent with a very high level of completeness—rare transmission failures spread across three fourths of participating hospitals—or very low completeness, with persistent failures at the majority of participating facilities. It’s not clear why this figure (26%) was calculated for the first four months of operation only.

p9, top paragraph notes that “Especially for ILI cases, a summer flu peak in 2004 was identified...” Figures 4 and 5 do not show a peak in ILI syndrome during the summer. Fever syndrome did peak during the summer, however, in the absence of laboratory confirmation of circulating influenza, and perhaps due to my ignorance of influenza patterns in Taiwan, it seems speculative to suggest that this summertime increase in ED visits had anything to do with influenza vaccinations not being available until after the summer season.

Table 2 should be deleted. The ED volume data presented are not that informative from a disease
surveillance perspective. Consider presenting information about coverage. It also is not clear what the Data Upload Days and Rate refer to. Is it the number and proportion of days when any data were received? Also, the wide range of daily visits (e.g. 1 to 20,456) indicate that days when transmission completely failed are included in the calculations, which would skew the means and increase the standard deviation. And how does a range of daily visits that includes 1 reconcile with a data upload rate of 100%?

Table 3: The significance of the age distribution of total ED visits presented in this table is unclear because the paper does not indicate how representative participating EDs are of all EDs in Taiwan. Readers may be more interested in seeing the distribution of key syndromes by age-group.

The “enterovirus-like illness” syndrome is unfamiliar and has an interesting seasonal pattern with two peaks, one in May-June and one in Sep-Nov. The authors should explain which ICD9 codes were included in this syndrome and comment on whether the pattern is consistent with the epidemiology of enterovirus infections in Taiwan.

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

The manuscript needs some editing. A few examples:

- p3 redundancy in first paragraph of Materials and methods.
- p5 “the MS-SQL server received on average the data of 7,378+/-2,893 data ED patients per day”
- p7 redundancy in first paragraph of Discussion
- p3 "Outbreaks of emerging infectious diseases are frequently first detected in the emergency healthcare setting." and again 7 lines down "many outbreaks of emerging infectious diseases require emergency health care."

p6 Details are needed to explain how a “computer program with data quality check-up procedures for automatic data quality control” helps hospitals and public health agencies maintain data quality. What, specifically, does the program look for. Is it fully ‘automatic’ or does a person review the results? If problems are found how are they corrected?

p8 Either more details are needed about the “feasible and easily installed feedback mechanism” or this reference should be deleted.

p7 Last paragraph of results “All these findings….” should be in the discussion.

-------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

The authors note that because only 25% of chief complaints are in English, they relied on ICD9 codes to categorize visits into syndromes. What were the barriers to categorizing Chinese chief complaints into syndromes?

Figure 2: It is not clear what the percents displayed on the map refer to, probably coverage but that should be indicated. They can be rounded to the nearest whole percent.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests