Reviewer's report

**Title:** Early detection of influenza outbreaks using the DC Department of Health's syndromic surveillance system

**Version:** 1  **Date:** 28 September 2009

**Reviewer:** Michael Jackson

**Reviewer's report:**

Early detection of influenza outbreaks using the DC Department of Health’s syndromic surveillance system

This manuscript is a revision of two manuscripts previously submitted to BMC Public Health; the two manuscripts have been combined into a single paper. This manuscript describes two sets of analyses. The first set of analyses attempt to optimize various aberration detection algorithms for detecting possible outbreaks in DC’s syndromic surveillance system. The second set compares several different data streams and syndrome definitions using emergency department (ED) data for early detection of seasonal influenza. The authors found that aberration detection methods are very sensitive to parameter choice, and that the optimal parameters may vary by data source. They also found that monitoring “unspecific infection” from ED visits to a children’s hospital provided the most timely detection of the start of influenza season, compared to other combinations of syndromes and emergency departments.

Major compulsory revisions

None.

Minor essential revisions

1. Combining the two former papers into one forces some information to be relegated to appendices or to supplementary on-line material, as the authors have done. However, the main manuscript should ideally be a stand-alone document. Overall the authors have done an admirable job with this. However, one place where information should be moved from the appendix to the main paper (or vice versa) is the description of the simulated outbreaks. The appendix describes how the simulated outbreaks were generated (using x added standardized counts). However, there are places in the main text and in the figures (page 13, second paragraph and figure 1, for example) where the authors refer to outbreaks of size x. This becomes confusing for readers who have not read the appendix before the results. For clarity, it would help to either move the description of the outbreak simulations to the main text, or to move discussion of outbreak size to the appendix or other supplementary material.

2. On page 15 at the top, the authors say "we recommend using (k,l) = (0.25,0.20) since it does outperform (k,l) = (0.25, 0.40) even if slightly."
Recommending specific values for parameters to be used in other contexts conflicts with the discussion: “The optimal parameters and detection algorithms that we found in this analysis, of course, apply only to the particular data sets analyzed…the results cannot be generalized to other data streams…” (page 21, third paragraph).

(3) An additional limitation that should be mentioned is that setting the false positive rate assumes there were no true outbreaks during the non-flu seasons. If there were outbreaks, this would cause the false positive rate to be overestimated.

(4) On page 17, the authors use CUSUM and EXPO to confirm that the “unspecified infection” is the most sensitive for detecting the onset of seasonal influenza. Why was MV CUSUM not also used here?

(5) On page 12, middle of the page, the following sentence has errors and is incomplete: “Initially, we examined how well each of the eight syndrome categories did along [sic] at flagging the beginning of the flu season depending.”

Discretionary revisions

(1) Figure 1B is apparently for the unspecified infection syndrome based on text and figure title in the manuscript, but the title in the .ppt file with the figures suggests the data come from the gastro syndrome.

(2) In the second paragraph of the introduction, a number of studies are summarized in which the performance of syndromic surveillance is presented as an improvement over other systems. A number of comparative statements are made, in which the comparison is not made clear. For example, “…respiratory syndromes represent an early indicator of impending influenza morbidity and mortality, sometimes by as much as three weeks” (page 4, second paragraph). Does this mean three weeks in advance of the actual influenza season, in advance of other surveillance methods, or something else? A similar issue exists for the description of the study by Olson et al.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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.