Reviewer's report

Title: Evaluation of school absenteeism data for early outbreak detection, New York City.

Version: 1 Date: 4 July 2005

Reviewer: Howard S Burkom

Reviewer's report:

General
Given the keen interest in evaluating possible sources of evidence for biosurveillance, it is important to have data source evaluations based on documented outbreaks as in the current report. However, these evaluations have admitted limitations that weaken any conclusions that may be drawn regarding the value of absenteeism data.

The papers abstract contains the statement that spatial analysis identified too many clusters to be useful. What the results actually showed was that the chosen implementation of spatial scan statistics is not practical for prospective monitoring with this data; leaping to a broad conclusion about spatial analysis is not justified.

These absenteeism data sets have limited usefulness for 2 reasons: (a) They are very nonspecific and only partially linked with illness. (b) The number of useful data days is severely limited by the school calendar, especially considering that about 20% of scheduled days should be ignored because they are close to holidays. From the discussion beginning Despite the removal of outliers, the authors imply that even more days should be removed to get purely calendar-related absences out of the data. The conclusion I would state is that much detailed analysis involving modeling and filtering of the data would be required to reduce the noise and, given (a) and (b), this analysis would be hard to justify for a health department with limited resources.

Any sweeping statement about the overall utility of absenteeism would require investigation of the hypothesis, stated in the full-text conclusion, that this data source can give earlier signals than a source based on acute clinical care, and such an investigation is impossible unless further modeling and data filtering can remove some of the noise.

The recent PLoS Medicine paper of Martin Kulldorff, with some of the same coauthors as the present work, has ideas that could be brought to bear. First, with 1160 participating schools, many with chronically high absenteeism or erratic reporting could be removed from the analysis and still preserve spatial coverage. The PLoS paper recommends a similar strategy to cope with missing data in some cases, and this strategy has avoided some excess clustering in ESSENCE absenteeism data. Second, the removal of space-by-day-of-week interaction could be relevant to this data source. Indeed, many noise-cancelling approaches based on detailed knowledge of the school system might be tried if there was evidence that the surveillance value of the data is worth the effort. These approaches could improve the temporal, city-wide results as well as the spatiotemporal ones.

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

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

In the abstract and full-text conclusions, avoid unjustified claims and point out that extensive modeling is required for enough noise reduction to establish the potential utility of absenteeism data in this school system.

p. 3: The authors correctly note that findings from one data stream can support or contradict. This
Remark should temper the previous statement. They might say that since findings from one data stream can support or contradict findings from another, multiple systems may increase or decrease detection performance, which is one good reason why their health department is not using algorithm results from absenteeism data.

P. 5-6: The CuSum reset criterion is nonstandard and did not eliminate multiday signals due to a spike on one day that was extreme enough to cause signals for 2 or 3 consecutive days. This is a weak condition that will let very high rates push later ones over the threshold until the outbreak has nearly fully subsided. By contrast, CDC's C3 CuSum algorithm does not add any sums from previous days that have flagged. The problem with the weaker condition is that if the system keeps alarming on a long downslope, sensitivity to secondary events is lost. Please just modify the explanation of this criterion.

P. 9: The final sentence in the paragraph about school A is still, 59 schools involved in clusters had single-day excesses of 200. I've read the paragraph several times and still don't see why that sentence is there. Please clarify others may not get it either.

P. 10: The statement Comparing these data to influenza surveillance is a good way is vacuous, and the authors have a valid point to make. A more descriptive phrasing such as: Trying to correlate peaks in daily time series from these data with documented rises in influenza test positives is an evidenced-based approach for testing the effectiveness would make this point.

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

P. 3: Correct the spelling of complementary.

P. 4: Reword for clarity: We calculated the median daily absentee rates among elementary/middle schools and separately among high schools and examined whether the absentee rates differed between these two groups.

P. 5: Refer to the Wilcoxon signed-rank test.

The clause we analyzed daily data retrospectively should be explained for those unused to this methodology. Make clear that the analysis was done to reflect the results of prospective monitoring (without knowledge of future data).

P. 6: Reword for clarity: "To assess geographic clustering we obtained from the DOE a more detailed dataset consisting of".

Discretionary revisions (which the author can choose to ignore)

P. 3: On my first reading of the statement Each of these syndromic systems relies on non-specific, pre-diagnostic data, I was tempted to object and replace systems with data sources. Though they are certainly different systems on the front end, the data sources may be treated as a single system for decision-making purposes. But this is a matter of bias, and it has not been proved that data fusion improves public health surveillance capability. However, in a unified system, the decision given in the papers final statement that prospective surveillance of school absenteeism will not be implemented in New York City might be different because the absentee data could be used to corroborate or add further information to indications from other data sources. For this paper, the authors should just be deliberate in their choice of systems or data sources in that statement.

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 importance in its field.
Quality of written English: Needs some language corrections before being published

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests.