Reviewer's report

Title: Using automated medical records for rapid identification of illness syndromes: the example of lower respiratory infection

Authors:

Ross Lazarus (ross.lazarus@channing.harvard.edu)
Kenneth P Kleinman (ken_kleinman@harvardpilgrim.org)
Inna Dashevsky (inna_dashevsky@harvardpilgrim.org)
Alfred DeMaria Jr. (alfred.demaria@state.ma.us)
Richard Platt (richard.platt@channing.harvard.edu)

Version: 4 Date: 9 Oct 2001

Reviewer: Dr James J. Cimino

Level of interest: A paper whose findings are important to those with closely related research interests

Advice on publication: Accept after revision, which I do not need to see

This is a well-organized, well-written paper on an interesting, timely topic. I have only minor suggestions to improve the paper:
- The Introduction could be improved by describing other work on surveillance systems. Two examples are mentioned in the Discussion (Armstrong and de Wit), but no description is provided.
- Also in the Introduction - has there been any other experience (published or unpublished) with the ESSENCE project?
- Somewhere, probably in the Discussion, you should note that others have found problems with ICD9 (such as Goldstein LB. Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: effect of modifier codes. Stroke. 29(8):1602-4, 1998 Aug) and defend your use of these data in that light.
- Penultimate paragraph in Methods, last sentence, "data...is" should be "data ... are"
- Discussion, 4th paragraph, change "ca" to "can"
- Same sentence: why not less than 24 hours? Why not hourly?
- Two paragraphs later: might not emergency room visits be more relevant than other ambulatory visits since they are, for the most part, unscheduled? They will tend to have more emergencies, which would be expected in a bioterrorism attack...
- Later.... you could also study a change in pattern with respect to repeat encounters (for example, untreated diseases would tend to increase the number and decrease the interval, compared to the usual pattern)
- Third-to-last paragraph of Discussion - "[cough] did not result in so many signals", yet it accounted for 52.8%
of the diagnoses. Something is confusing here...
- Same paragraph - improving the definitions might be helpful but how would you accomplish this within the ICD9 coding mechanism? Look for "fever" codes? Change ICD?

**Competing interests:**

None declared.