Reviewer's report

Title: Validity and timeliness of syndromic influenza surveillance during the autumn/winter wave of A(H1N1) influenza 2009 Results of emergency medical dispatch, ambulance and emergency department data from three European regions

Version: 1 Date: 21 January 2013

Reviewer: Kim Kavanagh

Reviewer's report:

Review of “Validity of timeliness of influenza syndromic surveillance during the autumn/winter wave of A(H1N1) influenza 2009”

1. Is the question posed by the authors well defined?
   Yes

2. Are the methods appropriate and well described?
   The methods used are generally appropriate however I believe there are a couple of sections to be amended and described better as detailed in the major compulsory revisions

3. Are the data sound?
   The data appear to be sound.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   Yes

5. Are the discussion and conclusions well balanced and adequately supported by the data?
   Yes but could be developed in line with comment below.

6. Are limitations of the work clearly stated?
   Limitations of the work are stated but could be added to in accordance with the comments I have raised.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
   Yes

8. Do the title and abstract accurately convey what has been found?
   Yes

9. Is the writing acceptable?
   Yes generally fine. I have highlighted a few small things to correct in my comments.
Major Compulsory Revisions

1. I do not believe that the assessment of correlation is appropriate for the data. Pearson correlation assumes normality of the data and a linear relationship and it is not apparent if that is the case here. A more appropriate technique would be to use spearman rank correlation or the cross correlation function (CCF). The CCF is likely to be most powerful when considering the degree of similarity between the reference data and emergency data time series. Please see and refer to the approaches summarised in Dailey et al. (2007) Timeliness of data sources used for influenza surveillance. J Am Med Inform Assoc 14:626-631

2. Dailey et al. also recommend running the reference data through the detection algorithm to see where this data would generate alarms. I think that doing so would make your conclusions on the timeliness of the system stronger.

3. The baseline period used in the CUSUM is defined for each data source but there is no assessment of the data which is observed in this period. This should be included in the paper. Was the process in control during the baseline period? In particular are the baseline periods for the Spanish and the Belgian data stable? They use a period when the first wave of the pandemic was occurring. This will surely be very different from the data in the Austrian baseline which is pre-pandemic. This should be added to the paper.

4. In an extension to point 2, there is no outline of the criteria used to signal alarms in the data. Assuming estimation of the mean count levels from the (hopefully) stable baseline period what technique was used to calculate the upper bounds of the CUSUM? Did this differ between the different data sources? This should be added to the methods and considered in the results section.

5. Assessment of the sensitivity and specificity of the method is assessed against the occurrence of the outbreak in the area but it is not known if the individuals in the data truly had Influenza A H1N1 or not therefore it is not truly sensitivity and specificity. The limitations of using this type of data to measure sensitivity and specificity should be added to the discussion.

Minor Essential Revisions

6. The last sentence in the conclusions section of the abstract is very long and confusing. Please reword.

7. In the background section reference to the Scottish of analysis of help-line data could be added with reference [10]. [10] doesn’t cover all of the UK despite the title and the reference below explains the regression modelling approach in more depth.


8. Tyrolean sick leave data is used as a substitute for sentinel data. Do you have any idea of how representative this data is likely to be? Are all individuals in
Austria required to have health insurance or would this represent a particular demographic of society? The representativeness of this reference data should be discussed in the text.

9. The start of the outbreak in the Austrian data is taken to be the official reported start. A few details of how this was decided from reference [25] would be helpful. Also how was the end date decided?

10. On page 12, “stratified application of CUSUM algorithms” is mentioned. How was this implemented? More detail is required in the text.

11. On page 13, the daily approach sensitivity and specificity paragraph is very repetitive of the paragraph preceding it. Perhaps it could be summarised to say that a similar approach was taken to with the weekly definition?

12. The reference data column in Table 1 is not clear, in particular the reporting delay illustrations. There needs to be text to explain what the shaded regions mean. This could be added to the footnote and explained in the text of the methods.

13. Table 3 should be amended. The caption should explain that the statistics show the daily counts. With this is mind what are the Z and H rows showing? I assume these are the test statistics rather than a count. Given that the test statistics have no interpretation without knowing the comparative cut-off value I think it would be wiser to report the p-values here. I would also change the row names to describe what the test is assessing rather than just the test name i.e. something like “Differences between week days evaluated by Kruskal-Wallis test”.

14. In the footnote of the Table 3 I don’t think that “A=data availability period” is very clear. Does this mean all of the data or the baseline period i.e. the data before period B?

15. In Table 3, it would have been helpful to see mean counts broken down by day of the week to understand the day of the week effects in each of the data sources.

16. Is the day of the week effect a weekend vs. week day effect? If so it might have been better to test Mon-Fri grouped against Sat &Sun grouped counts.

17. It would be helpful if the caption for Figure 1 could have dates along with week numbers. It would then be easier to relate to the dates in Table 1.

18. In Figure 1e (ED_ES) did the official pandemic definition really not hold for one week during the period? This seems strange and should be mentioned in the description of the reference data in the text if so.

19. Table 4 caption -please include dates with the week numbers here.

20. In the discussion there are several mentioned of performance and in particular “moderate performance” and “weakest performance”. These are rather vague terms and perhaps it would be better to talk specifically in terms of the validity and timeliness.

21. When p-values are included in the text the actual p-value should be reported rather than p<0.05 etc.
22. In the discussion (page 20), it is mention that the algorithm could be specified and fine-tuned to increase performance. This is very vague. Please include some specifics of what could be done to improve the model.

23. On page 20, the text says “As has been shown in other studies further specification of the CUSUM algorithm…”. Please include reference to these studies in the text.

24. On page 21, the CUSUM approach is justified by its ease of application. I do not think that other methods are particularly hard to apply and indeed the regression algorithm discussed in Kavanagh et al. is deliberately kept simple in order to maintain flexible application to different time series. I think that the wording of this statement should be reconsidered. Future work could consider regression approaches.

25. Some English needs correcting.
   a. Page 6, 3rd paragraph: “syndromic influenza surveillance systems focus not on..” change to “syndromic influenza surveillance systems do not focus on..”
   b. Same paragraphs: “diagnostic findings made during examinations” remove plural - “diagnostic findings made during examination”
   c. Page 9, 1st paragraph: “As substitute documented sick leaves…” change to “As a substitute documented sick leave…”
   d. Page 11. “Cases to which respiratory syndrome of ILI was assigned were aggregated to the variables respiratory syndrome or ILI cases per week and per day for further analyses” change to “Cases to which respiratory syndrome of ILI was assigned were aggregated by week and by day for further analyses”
   e. Page 18, 3rd paragraph. Change “..study of Cowling et al who applied also the CUSUM….”, to “..study of Cowling et al who also applied the CUSUM…."

Discretionary Revisions

26. In Table 1, for Austria and Belgium there are two rows of data but the reference data column applies to both but it looks like it is split into two rows as well. This confused me initially. Perhaps the reference data column would be better as a separate table.

27. In Table 2, the ED-BE breakdown is not provided because there are lots of different free text combinations. It may be better just to have this as one row for the total count rather than a breakdown of NAs which does not look good in the table.

28. In the text the data for Spain is introduced as Santander (page 8) but in the following text is referred to as Cantabrian. This is slightly confusing and it would be best to be consistent throughout.

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

Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests