Reviewer’s report

Title: Accuracy of epidemiological inferences based on publicly available information: retrospective comparative analysis of line lists of human cases infected with influenza A(H7N9) in China

Version: 1
Date: 12 February 2014
Reviewer: Walter Haas

Reviewer’s report:

In their well written manuscript the authors present a systematic comparison of different line lists – all based on publicly available data – with the line list compiled by the Chinese Centre for Disease Control and Prevention (China CDC). The latter serves as "gold standard" containing detailed epidemiologic information on each laboratory confirmed case of influenza A(H7N9). This analysis points out the important role of line list to allow the identification and evaluation of critical epidemiological parameters, especially in the beginning of the spread of a novel pathogen.

Major Compulsory Revisions

Abstract

1. P3, l16: “We characterized clinical severity and transmissibility…” The term “transmissibility” might be misleading, as the authors do not present data on R0 or risk of exposure, but rather did use the change in the incidence rates to demonstrate the impact of closure of poultry markets (as defined in the methods section, p9). Please rephrase.

2. P3, l20: “…, but there were more missing data on dates…” please include overall quantitative information about the extent of this information missing for the five line lists analyzed.

Background

3. The authors correctly mention that there is no official framework setup for early and rapid international exchange of epidemiological data. However, they should also address the role of WHO and the responsible national authorities.

Methods

4. P7, l18: “Each line list was compiled based on…” The Table S1 in the supplementary material is very much appreciated. The authors should include information about the case definition used to include cases as confirmed by the other line lists. As the publication delay and information between different data sources differs, usually for constructing a line list a number of decisions have to be made about which information should be included from which source. If there is information about the ranking of different data sources by the five line lists, this should also be included in Table S1.
5. P8, l3: Please specify which dates were selected for the estimation and explain the rationale for the selection of these dates. E.g. for the analysis of onset to hospitalization May 1st 2013 was chosen, even though the analysis included cases until May 31st.

6. P9, l2: Poisson regression models have been used. Had the data set been tested for overdispersion? (The cited Ref. #14 needs to be updated.) Was the model exactly applied as described in the reference? What adaptations were made when using the line list data from the other sources (e.g. taking into account the reporting delay)? Did the model control for differences in the age and sex distribution between the different line lists?

Results

7. P9, l19: “More than 90% of the cases could be matched...” What happened to the cases that could not be matched? Were these cases excluded from the analysis?

Discussion

8. Major limitations of publicly available information should be discussed. E.g. information published on individual cases from different sources cannot be matched without availability of a “gold standard”. Frequently there is the danger of duplicate reporting with differing dates, especially when family clusters are involved etc.

Minor Essential Revisions

9. P3, l22: the authors refer to HFR as “hospital fatality risk” should it rather read “hospital fatality ratio”? (also in the rest of the text)

10. P8, l1: Please refer to FIG S3 and additional information for explanation about how accuracy was defined and measured in this context.

11. P8, l15: it should read “…cumulative number of deaths by the cumulative number of hospitalized cases…”.

12. P12, l18: Add Ref. #27 as this also investigated severity based on case fatality ratios.

13. Please check and update reference list (e.g. missing information Ref. #9)

14. FIG2: Please add HFR1 and HFR2 in the label of the y-axis.

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests